Readings in the PHILOSOPHY OF LAW

Fifth Edition

John Arthur • William H. Shaw

Readings in the Philosophy of Law

Readings in the
Philosophy of Law

FIFTH EDITION

Readings in the Philosophy of Law

JOHN ARTHUR

Binghamton University,
State University of New York

WILLIAM H. SHAW

San Jose State University

Prentice Hall

Boston Columbus Indianapolis New York San Francisco Upper Saddle River
Amsterdam Cape Town Dubai London Madrid Milan Munich Paris Montreal Toronto
Delhi Mexico City Sao Paulo Sydney Hong Kong Seoul Singapore Taipei Tokyo

VP, Editorial Director: Leah Jewell
Editor in Chief: Dixon Musslewhite
Acquisitions Editor: Nancy Roberts
Editorial Assistant: Nart Varoqua
Marketing Assistant: Pat Walsh
Senior Marketing Assistant: Ashley Fallon
Production Manager: Fran Russello
Art Director: Jayne Conte
Cover Design: Axell Designs
Manager, Cover Visual Research & Permissions: Karen Sanatar
Cover Art: Getty Images Inc.
Full-Service Project Management: Shiny Rajesh/Integra Software Services, Ltd.
Printer/Binder: Courier Companies, Inc.
Text Font: 10/11 TimesTen Roman

Credits and acknowledgments for material borrowed from other sources and reproduced, with permission, in this textbook appear on appropriate page within text (or on pages 640–644).

Copyright © 2010, 2006, 2001, 1993, 1984, Pearson Education, Inc., publishing as Prentice Hall, 1 Lake Street, Upper Saddle River, NJ 07458. All rights reserved. Manufactured in the United States of America. This publication is protected by Copyright, and permission should be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or likewise. To obtain permission(s) to use material from this work, please submit a written request to Pearson Education, Inc., Permissions Department, Prentice Hall, 1 Lake Street, Upper Saddle River, NJ 07458.

Many of the designations by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this book, and the publisher was aware of a trademark claim, the designations have been printed in initial caps or all caps.

Library of Congress Cataloging-in-Publication Data

Readings in the philosophy of law/[edited by] John Arthur, William H. Shaw.—5th ed.
p. cm.
ISBN-13: 978-0-205-70809-3 (alk. paper)
ISBN-10: 0-205-70809-9 (alk. paper)
1. Law—Philosophy. 2. Law—United States. I. Arthur, John, 1946–2007. II. Shaw, William H., 1948–
K235.R39 2010
340'.1—dc22

2009040527

10 9 8 7 6 5 4 3

Prentice Hall
is an imprint of

PEARSON

www.pearsonhighered.com

ISBN 13: 978-0-205-70809-3
ISBN 10: 0-205-70809-9

CONTENTS

Preface *x*

PART I
LAW IN PRACTICE

1 ***The Adversary System and the Practice of Law*** *1*

Lawyers' Ethics in an Adversary System 1
Monroe H. Freedman

An Alternative to the Adversary System 10
John H. Langbein

Pure Legal Advocates and Moral Agents: Two Concepts of a Lawyer in an Adversary System 12
Elliot D. Cohen

2 ***The Rule of Law*** *22*

Rule of Law and the English Constitution 22
A. V. Dicey

Eight Ways to Fail to Make Law 27
Lon Fuller

The Rule of Law and Its Virtues 32
Joseph Raz

Grudge Informers and the Rule of Law 38
H. L. A. Hart

The Problem of the Grudge Informer 40
Lon Fuller

Detention of Enemy Combatants 44
Hamdi v. Rumsfeld

3 ***The Moral Force of Law*** *54*

Crito 54
Plato

The Justification of Civil Disobedience 59
John Rawls

On Not Prosecuting Civil Disobedience 67
Ronald Dworkin

PART II
LEGAL REASONING

4 ***Statutes*** *73*

On Interpretation: The Adultery Clause of the Ten Commandments 73
Sanford Levinson

Interpreting the Small Bird Act 75
Regina v. Ojibway

A Case Study in Interpretation: The Mann Act 77
Edward H. Levi

Cases Interpreting the Mann Act 81
Caminetti v. United States; Mortensen v. United States; and Cleveland v. United States

Interpretation 87
Tony Honoré

Who Is a Drunk Driver? 92
Kentucky v. Whitt

What Is a Vegetable? 93
Nix v. Hedden

"Use" of a Firearm 94
Smith v. United States

Can a Murderer Inherit? 98
Riggs v. Palmer

5 Precedents 103

Stare Decisis: The Use of Precedent 103
C. Gordon Post

Constrained by Precedent 113
Larry Alexander

Precedent and Legitimacy 117
Planned Parenthood v. Casey

6 Constitutional Interpretation 123

"Vague" Constitutional Clauses 123
Ronald Dworkin

The Due Process Clause 126
Rochin v. California

Interpreting the Constitution 132
Antonin Scalia

A Response to Scalia 138
Ronald Dworkin

The Second Amendment 144
District of Columbia v. Heller

In Defense of Looseness:
The Supreme Court and Gun Control 151
Richard A. Posner

PART III
THE NATURE OF LAW

7 Natural Law and Legal Positivism: Classical Perspectives 158

Traditional Natural Law Theory 158
Brian Bix

Summa Theologica 163
Thomas Aquinas

Blackstone's Commentaries 168
William Blackstone

The Province of Jurisprudence Determined 173
John Austin

8 Formalism and Legal Realism 179

The System of Law 179
Joseph H. Beale

The Path of the Law 182
Oliver Wendell Holmes

Realism and the Law 186
Jerome Frank

9 Morality and the Law 194

Positivism and the Separation of Law and Morals 194
H. L. A. Hart

Law as the Union of Primary and Secondary Rules 202
H. L. A. Hart

The Model of Rules 208
Ronald Dworkin

Natural Law Revisited 218
Ronald Dworkin

The Dependence of Morality on Law 224
Tony Honoré

10 International Law 230

International Law 230
H. L. A. Hart

The African Slave Trade 235
United States v. La Jeune Eugenie; The Antelope; and The Amistad

International Law and Individual Responsibility 240
The International Military Tribunal at Nuremberg

11 Law and Economics 246

The Economic Approach to Law 246
Richard A. Posner

Law and Economics: An Analysis and Critique 252
Andrew Altman

PART IV
PHILOSOPHICAL ISSUES IN CRIMINAL LAW

12 The Justification of Punishment 261

Who Should Be Punished? 261
The Case of the Dog Provetie

The Utilitarian Theory of Criminal Punishment 262
Richard B. Brandt

Persons and Punishment 268
Herbert Morris

The Moral Education Theory of Punishment 275
Jean Hampton

Restitution: A New Paradigm of Criminal Justice 279
Randy E. Barnett

13 The Rights of Defendants 286

Criminal Justice and the Negotiated Plea 287
Kenneth Kipnis

Convicting the Innocent 294
James McCloskey

The Dirty Harry Problem 301
Carl B. Klockars

A Debate on the Exclusionary Rule 309
Malcolm Richard Wilkey and Stephen H. Sachs

Interrogation and the Right to Counsel 320
Miranda v. Arizona; Brewer v. Williams; and Rhode Island v. Innis

Entrapment 335
United States v. Tobias

14 Sentencing 342

Capital Punishment 342
Gregg v. Georgia

Electric Shock: The Fairest Punishment of All 349
Graeme Newman

"Three Strikes" 354
Rummel v. Estelle

Racial Bias in Sentencing 357
McCleskey v. Kemp

15 Criminal Responsibility 360

Survival on a Lifeboat 360
The Queen v. Dudley and Stephens

The Principles of Criminal Law 365
Richard B. Brandt

Intention 370
H. L. A. Hart

Attempting the Impossible: The Crime That Never Was 377
Leo Katz

Drunkenness 384
People v. Koerber

The Battered Woman Syndrome 388
State v. Leidholm

Rape, Consent, and Mens Rea 391
Regina v. Morgan

Rape, Force, and Consent 393
Susan Estrich

Is the Insanity Test Insane? 401
R. J. Gerber

What Is So Special About Mental Illness? 409
Joel Feinberg

Executing Mentally Retarded Murderers 416
Atkins v. Virginia

PART V
PHILOSOPHICAL ISSUES IN CIVIL LAW

16 *Compensating for Private Harms: The Law of Torts* 423

The Efficiency of the Common Law 424
Andrew Altman

Negligence 427
William Prosser

Economic Efficiency and the "Hand Formula" 434
United States v. Carroll Towing Company

Foreseeable Risk 435
Stone v. Bolton

Tort Liability and Corrective Justice 438
Jeffrie G. Murphy and Jules L. Coleman

Negligence and Due Care 446
Palsgraf v. Long Island Railroad Co.

Loss, Agency, and Responsibility for Outcomes 452
Stephen R. Perry

Liability Without Causation? 457
Summers v. Tice

Legal Causation and the Desert Traveler 460
Amy Shapiro

A Duty to Rescue? 471
Yania v. Bigan; Farwell v. Keaton; and McFall v. Shimp

Wrongful Life and Wrongful Birth 476
Berman v. Allan

17 *Private Ownership: The Law of Property* 483

Property 484
John Locke

Property and Sovereignty 489
Morris Raphael Cohen

Intellectual Property 495
Richard A. Posner

Property Acquisition 501
Haslem v. Lockwood

Regulation of Private Property 504
Pennsylvania Coal Co. v. Mahon

Eminent Domain 507
Kelo v. City of New London

18 *Private Agreements: The Law of Contract* 515

The Basis of Contract 516
Morris Raphael Cohen

Contract as Promise 523
Charles Fried

Unconscionable Contracts 527
Williams v. Walker-Thomas Furniture Co.

Surrogate Mother Contracts 529
In the Matter of Baby M

Employment Contracts 537
Lochner v. New York; Muller v. Oregon; Coppage v. Kansas; and West Coast Hotel Co. v. Parrish

PART VI
PHILOSOPHICAL ISSUES IN CONSTITUTIONAL LAW

19 Church and State 544

A Letter Concerning Toleration 544
John Locke

School Prayer 548
Engel v. Vitale

Religious Freedom and Public Education 552
Wisconsin v. Yoder

Secular Humanism and Religious Establishment 555
Smith v. Board of Education of Mobile

20 Personal Liberty and Privacy 559

On Liberty 559
John Stuart Mill

The Right to Privacy 566
Griswold v. Connecticut

Homosexuality and the Right to Privacy 570
Lawrence v. Texas

Same-Sex Marriage 577
Goodridge v. Massachusetts Department of Public Health

Judicial Activism and Gay Marriage: A Debate 584
Lawrence Tribe and Richard Parker

21 Freedom of Speech 589

Of the Liberty of Thought and Discussion 589
John Stuart Mill

Flag Burning 593
Texas v. Johnson

Nazi Marches 597
Village of Skokie v. National Socialist Party

Obscenity 600
Paris Adult Theatre v. Slaton

Free Speech at School 603
Morse v. Frederick

22 Equality 610

Separate But Equal 610
Plessy v. Ferguson

The Great School Desegregation Case 613
Brown v. Board of Education

Wartime Internment of Japanese Americans 615
Korematsu v. United States

Racial Equality and Affirmative Action 620
Ronald Dworkin

Affirmative Action in Universities 628
Grutter v. Bollinger and Gratz v. Bollinger

Gender Discrimination 635
Michael M. v. Sonoma County Superior Court

Appendix: The Bill of Rights and the Fourteenth Amendment 638

Acknowledgments 640

PREFACE

We continue to be gratified by the reception that *Readings in the Philosophy of Law* has received, and are pleased to offer this fifth edition, which we think further improves on the earlier ones. Reviewers, friends, and our own experiences with the last edition all confirmed that the basic format and core readings should remain unchanged. We continue to believe that legal philosophy is its own subject, best approached by paying attention to how the law is organized and to the particular issues it raises. This is therefore not a book in applied ethics, nor one in political philosophy; philosophy of law is its own subject with its own fascinating set of problems.

The book is designed so that each section stands on its own, enabling instructors to address the topics in any order they choose and to skip issues and even entire sections if they wish. Readers familiar with earlier editions will find much continuity in the themes and topics treated. Believing it to be the soundest organizational structure, we continue to follow the traditional divisions among different fields of law. Initial sections on law in practice, legal reasoning, and the nature of law are followed by discussions of criminal, tort, contract, property, and constitutional law. These make up the major subdivisions of the book.

We have found it useful to begin undergraduate courses in the philosophy of law with discussions of the adversary system, lawyers' ethics, what is meant by the rule of law, and whether there is an obligation to obey the law and, if so, why. Many philosophy of law students plan to attend law school, and these topics are a natural way to introduce them to the subject. The first section of Part I now includes an essay by Elliot Cohen on two models of what it means to be a lawyer, while the section on the rule of law contains two new readings: A. V. Dicey's brief but classic exposition of the rule of law and the English constitution, and *Hamdi v. Rumsfeld*, in which the Supreme Court struggles with the implications of habeas corpus and the rule of law in the case of an American citizen held incommunicado as an enemy combatant in the war in Afghanistan. With separate opinions from four justices, *Hamdi* offers a thought-provoking range of conflicting perspectives.

Part II, Legal Reasoning, now includes the 1993 Supreme Court case *Smith v. U.S.*, which illustrates well a variety of issues involved in statutory interpretation, and a valuable overview of three different models of precedent by law professor Larry Alexander. But the biggest addition to Part II is a new section on constitutional interpretation. Following a short excerpt on "vague" constitutional clauses by the legal philosopher Ronald Dworkin, the important 1952 case *Rochin v. California* shows the Supreme Court debating the interpretation of one of those clauses as it asks whether a state's criminal proceedings have violated "due process of law." This is followed by Supreme Court justice Antonin Scalia's presentation and defense of *originalism*—his preferred method of constitutional interpretation—and a critique of Scalia by Dworkin that illustrates his own, rival approach. Scalia's originalism is in full evidence in his majority opinion in *District of Columbia v. Heller*. Here the Supreme Court wrestles with the meaning of the Second Amendment—a topic which it has seldom addressed and, then, only briefly. The concluding essay, a critique of the *Heller* decision by the conservative federal judge and legal thinker Richard A. Posner, offers yet another

approach to constitutional interpretation, distinct from both Scalia's and Dworkin's.

Changes to Parts III, IV, and V have been relatively minor, although attention should be drawn to two new cases: *People v. Koerber* on mens rea, intent, and drunkenness (in the section on criminal responsibility), and *Kelo v. City of New London* on eminent domain (in the section on property). Part VI, however, on constitutional law has been reorganized and strengthened, among other ways, by the addition of *Plessy v. Fergueson* and *Brown v. Board of Education*, cases that are always well worth teaching. A selection from *Gratz v. Bollinger* now supplements the existing excerpt from its companion case, decided the same day, *Grutter v. Bollinger* (also on affirmative action at the University of Michigan). Finally, the section on free expression contains the 2007 "BONG HiTS 4 JESUS" case, *Morse v. Frederick*, in which the Supreme Court examines freedom of speech in the context of public schools—with separate opinions by five different justices!

While presenting the readings in a coherent and interrelated format, so that they often address one another directly, we have at the same time sought to provide a textbook with ample resources to allow a variety of different courses to be taught, in different ways, and to different student audiences. These editorial decisions grow out of our experience in teaching courses in philosophy of law at a several colleges and universities, both private and public, to undergraduates with varying degrees of philosophical background and academic preparedness.

Philosophy of law is not a subject that can be made simple, but it can be made more accessible and more engaging to undergraduates than it has been in the past. Accordingly, we have endeavored to provide readings that students will find stimulating, yet that do justice to the subtlety and complexity of the philosophical problems in this field. We provide short introductions before, and questions for review and discussion after, each selection. These should enable even beginning students to follow the major arguments, at least in broad outline. In addition, we have edited the essays and cases with great care and abridged some of the readings from the previous edition in order to keep the focus sharply on the most relevant and important issues, avoiding needless technicality, immaterial digressions, and recondite issues.

* * * * * * * * * * *

The death of John Arthur has deprived this book of a co-editor, the world of philosophy of an insightful writer and able scholar, his department and professional colleagues of an energetic and committed member of the university community, his students of an inspirational and much loved teacher—and me, of one of the dearest friends and finest human beings I have ever known. Assisted by John's widow, the talented lawyer Amy Shapiro, I have, with a sad heart, completed this edition without him. But John's vision, intelligence, and wit live on in this book.

William H. Shaw

PART ONE

Law in Practice

CHAPTER 1

The Adversary System and the Practice of Law

We begin this book with essays on the nature of the practice of law for two reasons. First, the topic is of great practical and moral interest: Many who study the philosophy of law will be thinking about becoming lawyers. The oft-heard criticisms of lawyers and of the practice of law will be all too familiar to them. We all know that lawyers often bear the brunt of jokes about the practice of law. They are also frequently accused of demonstrating a lack of integrity or a willingness to do what may seem to be immoral acts. Why, for instance, does society tolerate the confidentiality of the lawyer–client relationship when it means that useful information is being withheld from the police? Or why are courtroom lawyers permitted to make truthful people appear to be liars and liars appear to be truthful? Second, beyond these practical and moral issues, this section also begins our inquiry into some of the other important topics that arise throughout the text, including the nature of law, the rule of law, and the justification of excluding relevant evidence in certain cases.

Lawyers' Ethics in an Adversary System

Monroe H. Freedman

In this selection from his book *Lawyers' Ethics in an Adversary System,* Monroe H. Freedman, formerly professor of law and dean of the Hofstra University School of Law, examines the obligations of criminal defense lawyers in three difficult, morally troubling cases: whether to keep knowledge of a client's crime confidential, whether to allow a client to present perjured testimony, and whether to destroy a truthful witness through tough cross-examination. Freedman notes the conflicting obligations facing the conscientious attorney, but he defends zealous and aggressive advocacy in these cases as a necessary part of our adversarial criminal justice system.

WHERE THE BODIES ARE BURIED: THE ADVERSARY SYSTEM AND THE OBLIGATION OF CONFIDENTIALITY

In a recent case in Lake Pleasant, New York, a defendant in a murder case told his lawyers about two other people he had killed and where their bodies had been hidden. The lawyers went there, observed the bodies, and took photographs of them. They did not, however, inform the authorities about the bodies until several months later, when their client had confessed to those crimes. In addition to withholding the information from police and prosecutors, one of the attorneys denied information to one of the victim's parents, who came to him in the course of seeking his missing daughter.

There were interesting reactions to that dramatic event. Members of the public were generally shocked at the apparent callousness on the part of the lawyers, whose conduct was considered typical of an unhealthy lack of concern by lawyers with the public interest and with simple decency. That attitude was encouraged by public statements by the local prosecutor, who sought to indict the lawyers for failing to reveal knowledge of a crime and for failing to see that dead bodies were properly buried. In addition, the reactions of lawyers and law professors who were questioned by the press were ambivalent and confused, indicating that few members of the legal profession had given serious thought to the fundamental questions of administration of justice and of professional responsibility that were raised by the case.

One can certainly understand the sense of moral compulsion to assist the parents and to give the dignity of proper burial to the victims. What seems to be less readily understood—but which, to my mind, throws the moral balance in the other direction—is the obligation of the lawyers to their client and, in a larger sense, to a system of administering justice which is itself essential to maintaining human dignity. In short, not only did the two lawyers behave properly, but they would have committed a serious breach of professional responsibility if they had divulged the information contrary to their client's interest. The explanation of that answer takes us to the very nature of our system of criminal justice and, indeed, to the fundamentals of our system of government. . . .

A trial is, in part, a search for truth. Accordingly, those basic rights are most often characterized as procedural safeguards against error in the search for truth. Actually, however, a trial is far more than a search for truth, and the constitutional rights that are provided by our system of justice may well outweigh the truth-seeking value—a fact which is manifest when we consider that those rights and others guaranteed by the Constitution may well impede the search for truth rather than further it. What more effective way is there, for example, to expose a defendant's guilt than to require self-incrimination, at least to the extent of compelling the defendant to take the stand and respond to interrogation before the jury? The defendant, however, is presumed innocent; the burden is on the prosecution to prove guilt beyond a reasonable doubt, and even the guilty accused has an "absolute constitutional right" to remain silent and to put the government to its proof.

Thus, the defense lawyer's professional obligation may well be to advise the client to withhold the truth. As Justice Jackson said, "Any lawyer worth his salt will tell the suspect in no uncertain terms to make no statement to police under any circumstances." Similarly, the defense lawyer is obligated to prevent the introduction of evidence that may be wholly reliable, such as a murder weapon seized in violation of the Fourth Amendment, or a truth ful but involuntary confession. Justice White has observed that although law enforcement officials must be dedicated to using only truthful evidence, "defense counsel has no comparable obligation to ascertain or present the truth. Our system assigns him a different mission. . . . [W]e . . . insist that he defend his client whether he is innocent or guilty." . . .

Before we will permit the state to deprive any person of life, liberty, or property, we require that certain processes be duly followed which ensure regard for the dignity of the individual, irrespective of the impact of those processes upon the determination of truth.

By emphasizing that the adversary process has its foundations in respect for human dignity, even at the expense of the search for truth, I do not mean to deprecate the search for truth or to suggest that the adversary system is not concerned with it. On the contrary, truth is a basic value, and the adversary system is one of the most efficient

and fair methods designed for determining it. That system proceeds on the assumption that the best way to ascertain the truth is to present to an impartial judge or jury a confrontation between the proponents of conflicting views, assigning to each the task of marshalling and presenting the evidence in as thorough and persuasive a way as possible. The truth-seeking techniques used by the advocates on each side include investigation, pretrial discovery, cross-examination of opposing witnesses, and a marshalling of the evidence in summation. Thus, the judge or jury is given the strongest possible view of each side, and is put in the best possible position to make an accurate and fair judgment. Nevertheless, the point that I now emphasize is that in a society that honors the dignity of the individual, the high value that we assign to truth-seeking is not an absolute, but may on occasion be subordinated to even higher values.

The concept of a right to counsel is one of the most significant manifestations of our regard for the dignity of the individual. No person is required to stand alone against the awesome power of the People of New York or the Government of the United States of America. Rather, every criminal defendant is guaranteed an advocate—a "champion" against a "hostile world," the "single voice on which he must rely with confidence that his interests will be protected to the fullest extent consistent with the rules of procedure and the standards of professional conduct." In addition, the attorney serves in significant part to assure equality before the law. Thus, the lawyer has been referred to as "the equalizer," who "places each litigant as nearly as possible on an equal footing under the substantive and procedural law under which he is tried."

The lawyer can serve effectively as advocate, however, "only if he knows all that his client knows" concerning the facts of the case. Nor is the client ordinarily competent to evaluate the relevance or significance of particular facts. What may seem incriminating to the client, may actually be exculpatory. For example, one client was reluctant to tell her lawyer that her husband had attacked her with a knife, because it tended to confirm that she had in fact shot him (contrary to what she had at first maintained). Having been persuaded by her attorney's insistence upon complete and candid disclosure, she finally "confessed all"—which permitted the lawyer to defend her properly and successfully on grounds of self-defense.

Obviously, however, the client cannot be expected to reveal to the lawyer all information that is potentially relevant, including that which may well be incriminating, unless the client can be assured that the lawyer will maintain all such information in the strictest confidence. "The purposes and necessities of the relation between a client and his attorney" require "the fullest and freest disclosures" of the client's "objects, motives and acts." If the attorney were permitted to reveal such disclosures, it would be "not only a gross violation of a sacred trust upon his part," but it would "utterly destroy and prevent the usefulness and benefits to be derived from professional assistance." That "sacred trust" of confidentiality must "upon all occasions be inviolable," or else the client could not feel free "to repose [confidence] in the attorney to whom he resorts for legal advice and assistance." Destroy that confidence, and "a man would not venture to consult any skillful person, or would only dare to tell his counselor half his case." The result would be impairment of the "perfect freedom of consultation by client with attorney," which is "essential to the administration of justice." Accordingly, the new Code of Professional Responsibility provides that a lawyer shall not knowingly reveal a confidence or secret of the client, nor use a confidence or secret to the disadvantage of the client, or to the advantage of a third person, without the client's consent. . . .

That is not to say, of course, that the attorney is privileged to go beyond the needs of confidentiality imposed by the adversary system, and actively participate in concealment of evidence or obstruction of justice. For example, in the *Ryder* case, which arose in Virginia several years ago, the attorney removed from his client's safe deposit box a sawed-off shotgun and the money from a bank robbery and put them, for greater safety, into the lawyer's own safe deposit box. The attorney, quite properly, was suspended from practice for 18 months. (The penalty might well have been heavier, except for the fact that Ryder sought advice from senior members of the bench and bar, and apparently acted more in ignorance than in venality.) The important difference between the *Ryder* case and the one in

Lake Pleasant lies in the active role played by the attorney in *Ryder* to conceal evidence. There is no indication, for example, that the attorneys in Lake Pleasant attempted to hide the bodies more effectively. If they had done so, they would have gone beyond maintaining confidentiality and into active participation in the concealment of evidence.

The distinction should also be noted between the attorney's knowledge of a past crime (which is what we have been discussing so far) and knowledge of a crime to be committed in the future. Thus, a major exception to the strict rule of confidentiality is the "intention of his client to commit a crime, and information necessary to prevent the crime." Significantly, however, even in that exceptional circumstance, disclosure of the confidence is only permissible, not mandatory. Moreover, a footnote in the Code suggests that the exception is applicable only when the attorney knows "beyond a reasonable doubt" that a crime will be committed. There is little guidance as to how the lawyer is to exercise the discretion to report future crimes. At one extreme, it seems clear that the lawyer should reveal information necessary to save a life. On the other hand, as will be discussed [below], the lawyer should not reveal the intention of a client in a criminal case to commit perjury in his or her own defense.

It has been suggested that the information regarding the two bodies in the Lake Pleasant case was not relevant to the crime for which the defendant was being prosecuted, and that, therefore, that knowledge was outside the scope of confidentiality. That point lacks merit for three reasons. First, an unsophisticated lay person should not be required to anticipate which disclosures might fall outside the scope of confidentiality because of insufficient legal relevance. Second, the information in question might well have been highly relevant to the defense of insanity. Third, a lawyer has an obligation to merge other, unrelated crimes into the bargained plea, if it is possible to do so. Accordingly, the information about the other murders was clearly within the production of confidentiality. . . .

In summary, the Constitution has committed us to an adversary system for the administration of criminal justice. The essentially humanitarian reason for such a system is that it preserves the dignity of the individual, even though that may occasionally require significant frustration of the search for truth and the will of the state. An essential element of that system is the right to counsel, a right that would be meaningless if the defendant were not able to communicate freely and fully with the attorney.

In order to protect the communication—and, ultimately, the adversary system itself—we impose upon attorneys what has been called the "sacred trust" of confidentiality. It was pursuant to that high trust that the lawyers acted in Lake Pleasant, New York, when they refrained from divulging their knowledge of where the bodies were buried. . . .

PERJURY: THE CRIMINAL DEFENSE LAWYER'S TRILEMMA

Is it ever proper for a criminal defense lawyer to present perjured testimony? . . . That question cannot be answered properly without an appreciation of the fact that the attorney functions in an adversary system of criminal justice which . . . imposes special responsibilities upon the advocate.

First, the lawyer is required to determine "all relevant facts known to the accused," because "counsel cannot properly perform their duties without knowing the truth." The lawyer who is ignorant of any potentially relevant fact "incapacitates himself to serve his client effectively," because "an adequate defense cannot be framed if the lawyer does not know what is likely to develop at trial."*

Second, the lawyer must hold in strictest confidence the disclosures made by the client in the course of the professional relationship. "Nothing is more fundamental to the lawyer-client relationship than the establishment of trust and confidence." The "first duty" of an attorney is "to keep the secrets of his clients." If this were not so, the client would not feel free to confide fully, and the lawyer would not be able to fulfill the obligation to ascertain all relevant facts. Accordingly, defense counsel is required to establish "a relationship of trust and confidence" with the accused, to explain "the necessity of full disclosure of all facts," and to

**American Bar Association Canons of Professional Ethics*, 15.

explain to the client "the obligation of confidentiality which makes privileged the accused's disclosures."

Third, the lawyer is an officer of the court, and his or her conduct before the court "should be characterized by candor."

As soon as one begins to think about those responsibilities, it becomes apparent that the conscientious attorney is faced with what we may call a trilemma—that is, the lawyer is required to know everything, to keep it in confidence, and to reveal it to the court. Moreover, the difficulties presented by those conflicting obligations are particularly acute in the criminal defense area because of the presumption of innocence, the burden upon the state to prove its case beyond a reasonable doubt, and the right to put the prosecution to its proof.

Before addressing the issue of the criminal defense lawyer's responsibilities when the client indicates to the lawyer the intention to commit perjury in the future, we might note the somewhat less difficult question of what the lawyer should do when knowledge of the perjury comes after its commission rather than before it. Although there is some ambiguity in the most recent authorities, the rules appear to require that the criminal defense lawyer should urge the client to correct the perjury, but beyond that, the obligation of confidentiality precludes the lawyer from revealing the truth. . ...

If we recognize that professional responsibility requires that an advocate have full knowledge of every pertinent fact, then the lawyer must seek the truth from the client, not shun it. That means that the attorney will have to dig and pry and cajole, and, even then, the lawyer will not be successful without convincing the client that full disclosure to the lawyer will never result in prejudice to the client by any word or action of the attorney. That is particularly true in the case of the indigent defendant, who meets the lawyer for the first time in the cell block or the rotunda of the jail. The client did not choose the lawyer, who comes as a stranger sent by the judge, and who therefore appears to be part of the system that is attempting to punish the defendant. It is no easy task to persuade that client to talk freely without fear of harm. . . .

Assume the following situation. Your client has been falsely accused of a robbery committed at 16th and P Streets at 11:00 p.m. He tells you at first that at no time on the evening of the crime was he within six blocks of that location. However, you are able to persuade him that he must tell you the truth and that doing so will in no way prejudice him. He then reveals to you that he was at 15th and P Streets at 10:55 that evening, but that he was walking east, away from the scene of the crime, and that, by 11:00 p.m., he was six blocks away. At the trial, there are two prosecution witnesses. The first mistakenly, but with some degree of persuasiveness, identifies your client as the criminal. At that point the prosecution's case depends upon that single witness, who might or might not be believed. The second prosecution witness is an elderly woman who is somewhat nervous and who wears glasses. She testifies truthfully and accurately that she saw your client at 15th and P Streets at 10:55 p.m. She has corroborated the erroneous testimony of the first witness and made conviction extremely likely. However, on cross-examination her reliability is thrown into doubt through demonstration that she is easily confused and has poor eyesight. Thus, the corroboration has been eliminated, and doubt has been established in the minds of the jurors as to prosecution's entire case.

The client then insists upon taking the stand in his own defense, not only to deny the erroneous evidence identifying him as the criminal, but also to deny the truthful, but highly damaging, testimony of the corroborating witness who placed him one block away from the intersection five minutes prior to the crime. Of course, if he tells the truth and thus verifies the corroborating witness, the jury will be more inclined to accept the inaccurate testimony of the principal witness, who specifically identified him as the criminal.

In my opinion, the attorney's obligation in such a situation would be to advise the client that the proposed testimony is unlawful, but to proceed in the normal fashion in presenting the testimony and arguing the case to the jury if the client makes the decision to go forward. Any other course would be a betrayal of the assurances of confidentiality given by the attorney in order to induce the client to reveal everything, however damaging it might appear. . . .

For example, how would [one] resolve the following case? The prosecution witness testified that

the robbery had taken place at 10:15, and identified the defendant as the criminal. However, the defendant had a convincing alibi for 10:00 to 10:30. The attorney presented the alibi, and the client was acquitted. The alibi was truthful, but the attorney knew that the prosecution witness had been confused about the time, and that his client had in fact committed the crime at 10:45. (Ironically, that same attorney considers it clearly unethical for a lawyer to present the false testimony on behalf of the innocent defendant in the case of the robbery at 16th and P Streets.) Should the lawyer have refused to present the honest alibi? How could he possibly have avoided doing so? Was he contributing to wise and informed judgment when he did present it?

The most obvious way to avoid the ethical difficulty is for the lawyer to withdraw from the case, at least if there is sufficient time before trial for the client to retain another attorney. The client will then go to the nearest law office, realizing that the obligation of confidentiality is not what it has been represented to be, and withhold incriminating information or the fact of guilt from the new attorney. In terms of professional ethics, the practice of withdrawing from a case under such circumstances is difficult to defend, since the identical perjured testimony will ultimately be presented. Moreover, the new attorney will be ignorant of the perjury and therefore will be in no position to attempt to discourage the client from presenting it. Only the original attorney, who knows the truth, has that opportunity, but loses it in the very act of evading the ethical problem.

The difficulty is all the more severe when the client is indigent. In that event, the client cannot retain other counsel, and in many jurisdictions it is impossible for appointed counsel or a public defender to withdraw from a case except for extraordinary reasons. Thus, the attorney can successfully withdraw only by revealing to the judge that the attorney has received knowledge of the client's guilt, or by giving the judge a false or misleading reason for moving for leave to withdraw. However, for the attorney to reveal knowledge of the client's guilt would be a gross violation of the obligation of confidentiality, particularly since it is entirely possible in many jurisdictions that the same judge who permits the attorney to withdraw will subsequently hear the case and sentence the defendant. Not only will the judge then have personal knowledge of the defendant's guilt before the trial begins, but it will be knowledge of which the newly appointed counsel for the defendant will very likely be ignorant.

Even where counsel is retained, withdrawal may not be a practical solution either because trial has begun or it is so close to trial that withdrawal would leave the client without counsel, or because the court for other reasons denies leave to withdraw. Judges are most reluctant to grant leave to withdraw during the trial or even shortly before it because of the power that that would give to defendants to delay the trial date or even to cause a series of mistrials. . . .

Since there are actually three obligations that create the difficulty—the third being the attorney's duty to learn all the facts—there is, of course, another way to resolve the difficulty. That is, by "selective ignorance." The attorney can make it clear to the client from the outset that the attorney does not want to hear an admission of guilt or incriminating information from the client. That view, however, puts an unreasonable burden on the unsophisticated client to select what to tell and what to hold back, and it can seriously impair the attorney's effectiveness in counselling the client and in trying the case.

For example, one leading attorney, who favors selective ignorance to avoid the trilemma, told me about one of his own cases in which the defendant assumed that the attorney would prefer to be ignorant of the fact that the defendant had been having sexual relations with the chief defense witness. As a result of the lawyer's ignorance of that fact, he was unable to minimize its impact by raising it with potential jurors during jury selection and by having the defendant and the defense witness admit it freely on direct examination. Instead, the first time the lawyer learned about the illicit sexual relationship was when the prosecutor dramatically obtained a reluctant admission from the defense witness on cross-examination. The defense attorney is convinced that the client was innocent of the robbery with which he had been charged, but the defendant was nevertheless found guilty by the jury—in the attorney's own opinion because the defendant was guilty of fornication, a far less serious offense for which he had not been charged.

The question remains: What should the lawyer do when faced with the client's insistence upon taking the stand and committing perjury? It is in response to that question that the Standards present a most extraordinary solution. If the lawyer knows that the client is going to commit perjury, Section 7.7 of the Standards requires that the lawyer "must confine his examination to identifying the witness as the defendant and permitting him to make his statement." That is, the lawyer "may not engage in direct examination of the defendant . . . in the conventional manner." Thus, the client's story will become part of the record, although without the attorney's assistance through direct examination. The general rule, of course, is that in closing argument to the jury "the lawyer may argue all reasonable inferences from the evidence in the record." Section 7.7 also provides, however, that the defense lawyer is forbidden to make any reference in closing argument to the client's testimony.

There are at least two critical flaws in that proposal. The first is purely practical: The prosecutor might well object to testimony from the defendant in narrative form rather than in the conventional manner, because it would give the prosecutor no opportunity to object to inadmissible evidence prior to the jury's hearing it. The Standards provide no guidance as to what the defense attorney should do if the objection is sustained.

More importantly, experienced trial attorneys have often noted that jurors assume that the defendant's lawyer knows the truth about the case, and that the jury will frequently judge the defendant by drawing inferences from the attorney's conduct in the case. There is, of course, only one inference that can be drawn if the defendant's own attorney turns his or her back on the defendant at the most critical point in the trial, and then, in closing argument, sums up the case with no reference to the fact that the defendant has given exculpatory testimony. . . .

It would appear that the ABA Standards have chosen to resolve the trilemma by maintaining the requirements of complete knowledge and of candor to the court, and sacrificing confidentiality. Interestingly, however, that may not in fact be the case. I say that because the Standards fail to answer a critically important question: Should the client be told about the obligation imposed by Section 7.7? That is, the Standards ignore the issue of whether the lawyer should say to the client at the outset of their relationship: "I think it's only fair that I warn you: If you should tell me anything incriminating and subsequently decide to deny the incriminating facts at trial, I would not be able to examine you in the ordinary manner or to argue your untrue testimony to the jury." The Canadian Bar Association, for example, takes an extremely hard line against the presentation of perjury by the client, but it also explicitly requires that the client be put on notice of that fact. Obviously, any other course would be a betrayal of the client's trust, since everything else said by the attorney in attempting to obtain complete information about the case would indicate to the client that no information thus obtained would be used to the client's disadvantage.

On the other hand, the inevitable result of the position taken by the Canadian Bar Association would be to caution the client not to be completely candid with the attorney. That, of course, returns us to resolving the trilemma by maintaining confidentiality and candor, but sacrificing complete knowledge—a solution which, as we have already seen, is denounced by the Standards as "unscrupulous," "most egregious," and "professional impropriety."

Thus, the Standards, by failing to face up to the question of whether to put the client on notice, take us out of the trilemma by one door only to lead us back by another. . . .

Taking into account, therefore, . . . the practical and constitutional difficulties encountered by any of the alternatives to strict maintenance of confidentiality, . . . I continue to stand with those lawyers who hold that "the lawyer's obligation of confidentiality does not permit him to disclose the facts he has learned from his client which form the basis for his conclusion that the client intends to perjure himself." What that means—necessarily, it seems to me—is that the criminal defense attorney, however unwillingly in terms of personal morality, has a professional responsibility as an advocate in an adversary system to examine the perjurious client in the ordinary way and to argue to the jury, as evidence in the case, the testimony presented by the defendant. . . .

CROSS-EXAMINATION: DESTROYING THE TRUTHFUL WITNESS

More difficult than the question of whether the criminal defense lawyer should present known perjury, is the question of whether the attorney should cross-examine a witness who is testifying accurately and truthfully, in order to make the witness appear to be mistaken or lying. The issue was raised effectively in a symposium on legal ethics through the following hypothetical case.

The accused is a drifter who sometimes works as a filling station attendant. He is charged with rape, a capital crime. You are his court-appointed defense counsel. The alleged victim is the twenty-two-year-old daughter of a local bank president. She is engaged to a promising young minister in town. [The drifter claims she consented.] . . .

You learn that the victim has had affairs with two local men from good families. Smith, one of these young men, admits that the victim and he went together for some time, but refuses to say whether he had sexual intercourse with her and indicates he has a low opinion of you for asking. The other, Jones, apparently a bitterly disappointed and jealous suitor, readily states that he frequently had intercourse with the victim, and describes her behavior toward strange men as scandalous. He once took her to a fraternity dance, he says, and, having noticed she had been gone for some time, discovered her upstairs with Smith, a fraternity brother, on a bed in a state of semi-undress. He appears eager to testify and he states that the girl got what she'd always been asking for. You believe Jones, but are somewhat repelled by the disappointed suitor's apparent willingness to smear the young woman's reputation.

Suppose the accused, after you press him, admits that he forced himself on the victim and admits that his first story was a lie. He refuses to plead guilty to the charge or any lesser charge. He says that he can get away with his story, because he did once before in California.

Should the defense lawyer use the information supplied by Jones to impeach the young woman and, if necessary, call Jones as a witness? . . .

That case takes us to the heart of my disagreement with the traditional approach to dealing with difficult questions of professional responsibility. That approach has two characteristics. First, in a rhetorical flourish, the profession is committed in general terms to all that is good and true. Then, specific questions are answered by uncritical reliance upon legalistic norms, regardless of the context in which the lawyer may be acting, and regardless of the motive and the consequences of the act. Perjury is wrong, and therefore no lawyer, in any circumstance, should knowingly present perjury. Cross-examination, however, is good, and therefore any lawyer, under any circumstances and regardless of the consequences, can properly impeach a witness through cross-examination. The system of professional responsibility that I have been advancing, on the other hand, is one that attempts to deal with ethical problems in context—that is, as part of a functional sociopolitical system concerned with the administration of justice in a free society—and giving due regard both to motive and to consequences. In that respect, the debate returns us to some fundamental philosophical questions that have not been adequately developed in the literature of professional responsibility. . . .

One of the major flaws in the traditional approach to legal ethics is that it seeks to an answer the difficult questions in a legalistic fashion at the personal level, but begs completely the critical questions raised at the systemic level. Thus, if you say to a lawyer: "Lawyers are under a moral duty not to participate in the presentation of perjury, and therefore you are required to act in a way contrary to your client's interest if the client insists upon committing perjury," the lawyer is entitled to respond: "Let us consider your maxim. If it is embodied into the system as a universal law to be applied to all lawyers in all circumstances, would the maxim destroy itself and be destructive of the system?"

As we have seen [previously], the system requires the attorney to know everything that the client knows that is relevant to the case. In order to enable the lawyer to obtain that information, the system provides for an obligation of confidentiality, designed to protect the client from being prejudiced by disclosures to the attorney. In addition, the attorney is required to impress upon the client the obligation of confidentiality in order to induce the client to confide freely and fully.

Let us return, then, to the case involving the street robbery at 16th and P Streets, in which the

defendant has been wrongly identified as the criminal, but correctly identified by the nervous, elderly woman who wears eyeglasses, as having been only a block away five minutes before the crime took place. If the woman is not cross-examined vigorously and her testimony shaken, it will serve to corroborate the erroneous evidence of guilt. On the other hand, the lawyer could take the position that since the woman is testifying truthfully and accurately, she should not be made to appear to be mistaken or lying. But if a similar course were to be adopted by every lawyer who learned the truth through confidential disclosures from the client, such disclosures would soon cease to be made. The result, for practical purposes, would be identical with the practice, disapproved in the ABA Standards, of "selective ignorance," in which the client is warned not to reveal to the lawyer anything that might prove embarrassing and prevent the lawyer from doing a vigorous job of presenting evidence and cross-examining. Of course, if that is the result we want, it would be far better that lawyers take a direct and honest approach with their clients, telling them to be less than candid, rather than lying to their clients by impressing upon them a bond of trust that the lawyers do not intend to maintain. Thus, when we examine the problem in a systemic context, we reach the conclusion . . . supporting cross-examination of the prosecutrix in the rape case.

Obviously, however, the rape case is a much harder one, because the injury done to the prosecutrix is far more severe than the more limited humiliation of the public-spirited and truthful witness in the case of the street robbery. In addition, in the rape case, the lawyer is acting pursuant to a manifestly irrational rule, that is, one that permits the defense to argue that the prosecutrix is the kind of person who would have sexual intercourse with a stranger because she has had sexual relations with two men whom she knew in wholly different social circumstances. Irrational or not, however, in those jurisdictions in which the defense of unchastity is still the law, the attorney is bound to provide it on the client's behalf. For the lawyer who finds the presentation of that defense, and perhaps others in rape cases, to go beyond what he or she can in good conscience do, there are two courses that should be followed. The first is to be active in efforts to reform the law in that regard; the second is to decline to accept the defense of rape cases, on the grounds of a conflict of interest (a strong personal view) that would interfere with providing the defendant with his constitutional rights to effective assistance of counsel.

REVIEW AND DISCUSSION QUESTIONS

1. What moral and professional conflict is raised by the Lake Pleasant, buried-bodies case? Explain how Freedman appeals to our adversarial system to defend confidentiality. What are the limits to what an attorney may do to assist a client? How does the *Ryder* case differ from the Lake Pleasant case?
2. What does Freedman mean by the "lawyer's trilemma"? Why does he criticize the treatment of the issue by the American Bar Association (ABA) Standards?
3. Should an attorney allow a client to present perjured testimony to the court? Is the Canadian Bar Association's rule an improvement over the system that Freedman defends?
4. How ought an attorney handle the cross-examination of the truthful witness?
5. Does Freedman's defense of the adversary system tip the balance too much in favor of criminal defendants?
6. Freedman discusses criminal cases, but our society also uses an adversary system for handling civil cases. To what extent can Freedman's arguments be used to defend zealous advocacy and aggressive legal tactics in civil cases?

An Alternative to the Adversary System

John H. Langbein

It is often assumed by practicing lawyers as well as the wider public that the adversary system is if not perfect, at least better than its alternatives. But why should that be assumed, especially in light of the fact that so many other nations use a different system? Indeed it has seemed to many that the adversary system makes demands on lawyers that cannot possibly be met, since it asks that they both represent their clients' interests and that they serve justice as "officers of the Court." In this essay, John H. Langbein describes the differences between the Anglo-American and German systems of civil procedure. Civil procedure (as opposed to criminal procedure) is the process by which the legal system resolves disputes between private parties, for example, disputes over the terms of a contract, damage caused by negligence, property transactions, marriage, or corporate and business relationships—in other words, much of the law. As Langbein points out, the major differences between the adversary system and what is sometimes called an "inquisitorial" one are first that the Court plays the dominant role in gathering the evidence, rather than the lawyers, and second that the trial is not treated as a separate, brief period in which the lawyers and clients present their cases to the judge or jury but rather as an extended process during which evidence is gathered and assessed.

There are two fundamental differences between German and Anglo-American civil procedure, and these differences lead in turn to many others. First, the court rather than the parties' lawyers takes the main responsibility for gathering and sifting evidence, although the lawyers exercise a watchful eye over the court's work. Second, there is no distinction between pretrial and trial, between discovering evidence and presenting it. Trial is not a single continuous event. Rather, the court gathers and evaluates evidence over a series of hearings, as many as the circumstances require. . . .

From the standpoint of comparative civil procedure, the most important consequence of having judges direct fact-gathering in this episodic fashion is that . . . in German procedure the court ranges over the entire case, constantly looking for the jugular—for the issue of law or fact that might dispose of the case. Free of constraints that arise from party presentation of evidence, the court investigates the dispute in the fashion most likely to narrow the inquiry. A major job of counsel is to guide the search by directing the court's attention to particularly cogent lines of inquiry. . . .

The episodic character of German civil procedure—Benjamin Kaplan called it the "conference method" of adjudication—has other virtues: It lessens tension and theatrics, and it encourages settlement. Countless novels, movies, plays, and broadcast serials attest to the dramatic potential of the Anglo-American trial. . . . German civil proceedings have the tone not of the theatre, but of a routine business meeting—serious rather than tense. When the court inquires and directs, it sets no stage for advocates to perform. The forensic skills of counsel can wrest no material advantage. . . .

In this business-like system of civil procedure the tradition is strong that the court promotes compromise. The judge who gathers the facts soon knows the case as well as the litigants do, and he concentrates each subsequent increment of fact-gathering on the most important issues still unresolved. As the case progresses the judge discusses it with the litigants, sometimes indicating provisional views of the likely outcome. He is, therefore, strongly positioned to encourage a litigant to abandon a case that is turning out to be weak or hopeless, or to recommend settlement. . . .

Adversary control of fact-gathering in [U.S.] procedure entails a high level of conflict between partisan advantage and orderly disclosure of the relevant information. Marvin Frankel put this point crisply when he said that "it is the rare case in which either side yearns to have the witnesses, or anyone, give *the whole truth.*"

If we had deliberately set out to find a means of impairing the reliability of witness testimony, we

could not have done much better than the existing system of having partisans prepare witnesses in advance of trial and examine and cross-examine them at trial. . . . "[T]he partisan nature of trials tends to make partisans of the witnesses."

Cross-examination at trial—our only substantial safeguard against this systematic bias in the testimony that reaches our courts—is a frail and fitful palliative. Cross-examination is too often ineffective to undo the consequences of skillful coaching. Further, because cross-examination allows so much latitude for bullying and other truth-defeating stratagems, it is frequently the source of fresh distortion when brought to bear against truthful testimony. As a leading litigator boasted recently in an ABA publication: "By a carefully planned and executed cross-examination, I can raise at least a slight question about the accuracy of [an adverse] witness's story, or question his motives or impartiality."

When we cross the border into German civil procedure, we leave behind all traces of this system of partisan preparation, examination, and cross-examination of witnesses. German law distinguishes parties from witnesses. A German lawyer must necessarily discuss the facts with his client, and based on what his client tells him and on what the documentary record discloses, the lawyer will nominate witnesses whose testimony might turn out to be helpful to his client. As the proofs come in, they may reveal to the lawyer the need to nominate further witnesses for the court to examine. But the lawyer stops at nominating; virtually never will he have occasion for out-of-court contact with a witness. Not only would such contact be a serious ethical breach, it would be self-defeating. "German judges are given to marked and explicit doubts about the reliability of the testimony of witnesses who previously have discussed the case with counsel or who have consorted unduly with a party." . . .

Equality of representation. The German system gives us a good perspective on another great defect of adversary theory, the problem that the Germans call "Waffenungleichheit"—literally, inequality of weapons, or in this instance, inequality of counsel. In a fair fight the pugilists must be well matched. . . . The simple truth is that very little in our adversary system is designed to match combatants of comparable prowess, even though adversarial prowess is a main factor affecting the outcome of litigation. Adversary theory thus presupposes a condition that adversary practice achieves only indifferently. It is a rare litigator in the United States who has not witnessed the spectacle of a bumbling adversary whose poor discovery work or inability to present evidence at trial caused his client to lose a case that should have been won. Disparity in the quality of legal representation can make a difference in Germany, too, but the active role of the judge places major limits on the extent of the injury that bad lawyering can work on a litigant. In German procedure both parties get the same fact-gatherer—the judge. . . .

Prejudgment. Perhaps the most influential justification for adversary domination of fact-gathering is that nonadversarial procedure risks prejudgment—that is, prematurity in judgment. . . .

In German procedure counsel oversees and has means to prompt a flagging judicial inquiry; but quite apart from that protection, is it really true that a "familiar pattern" would otherwise beguile the judge into investigating too sparingly? If so, it seems odd that this asserted "natural human tendency" towards premature judgment does not show up in ordinary business and personal decision-making, whose patterns of inquiry resemble the fact-gathering process in German civil procedure. Since the decision-maker does his own investigating in most of life's decisions, it seems odd to despair of prematurity only when that normal mode of decision-making is found to operate in a courtroom. . . .

Depth. Concern about prematurity shades into a different issue: how to achieve appropriate levels of depth in fact-gathering. Extra investment in search can almost always turn up further proofs that would be at least tenuously related to the case. Adversary domination of fact-gathering privatizes the decision about what level of resources to invest in the case. . . .

[German fact-gathering] does indeed contrast markedly with the inclination of American litigators "to leave no stone unturned, provided, of course, they can charge by the stone." The primary reason that German courts do less fact-gathering than American lawyers is that the Germans eliminate the waste. . . .

Because German procedure places upon the judge the responsibility for fact-gathering, the danger arises that the job will not be done well. The American system of partisan fact-gathering has the virtue of its vices: It aligns responsibility with incentive. Each side gathers and presents proofs according to its own calculation of self-interest. This privatization is an undoubted safeguard against official sloth. After all, who among us has not been treated shabbily by some lazy bureaucrat in a government department? And who would want to have that ugly character in charge of one's lawsuit?

The answer to that concern in the German tradition is straightforward: The judicial career must be designed in a fashion that creates incentives for diligence and excellence. The idea is to attract very able people to the bench, and to make their path of career advancement congruent with the legitimate interests of the litigants.

REVIEW AND DISCUSSION QUESTIONS

1. Describe the fundamental differences between the German inquisitorial system and the Anglo-Saxon adversary system.
2. Why has the inquisitorial system been described as a "conference method" of deciding disputes?
3. Describe the advantages that Langbein believes the inquisitorial system has, as compared with the adversarial system.
4. What dangers does Langbein see in the inquisitorial system? How can they best be addressed?
5. Is the case for the adversary system stronger for criminal law than civil law? Explain.
6. How might Monroe Friedman respond to Langbein?

Pure Legal Advocates and Moral Agents: Two Concepts of a Lawyer in an Adversary System

Elliot D. Cohen

Can an effective lawyer also be a morally good person? In addressing that question, Elliot Cohen first identifies the characteristics that mark a morally good person. He then describes the *pure legal advocate concept* of a lawyer and argues that a pure legal advocate will fall far short of being a morally good person. Cohen then goes on to explain the *moral agent concept* of a lawyer and to spell out the principles that a morally good lawyer should try to conform to. These principles are not unconditional, however, and there is no simple formula for resolving possible conflicts between them.

It is sometimes asked whether a good lawyer in an adversary system can also be a good person. We must first notice that there are two different senses of the term *good* employed in this question. In its first occurrence, *good* may be taken in its instrumental sense to mean, roughly, *effective.* In its second occurrence, *good* may be taken in its moral sense to mean *morally good.* Thus the question is whether an effective lawyer can also be a morally good person. And the latter question, it is clear, can be answered only if we have some idea of what we mean by a morally good person, and by an effective lawyer.

Accordingly, in this paper I shall first outline what we take to be salient marks of a morally good person. Second, I shall examine one sense of a

lawyer, what we call the *pure legal advocate concept*, in which a good lawyer does *not* satisfy our criteria of a morally good person. Third, I shall examine a further concept of a lawyer, what might be called the *moral agent concept*, according to which a good lawyer is, *ipso facto*, a morally good person.

MORALLY GOOD PERSONS

Following one tradition, let us say that a morally good person is a person who, through exercise and training, has cultivated certain morally desirable traits of character; the latter traits being constituted by dispositions to act, think, and feel in certain ways, under certain conditions, which are *themselves* morally desirable. What traits of character in particular are morally desirable and to what extent and in what combinations they must be cultivated in order for a person to be morally good are admittedly no settled matters. Still, there are some traits which at least most of us would countenance as being important, if not essential, ingredients of the morally good personality. It is such traits of character with which I shall be concerned, particularly those among them which seem to be the most relevant to legal practice.

What then are some such characteristic marks of a morally good person?

1. We would not ordinarily countenance a person as being morally good if we believed that he was not a *just* person, that is, if we thought that he was not disposed toward treating others justly. There are, however, two senses of *just* and *unjust* in which a person may be said to treat others justly of unjustly.

First, a person may be said to treat others justly when, in distributing some good or service among them, she observes the principle of treating relevantly similar cases in a similar fashion; and she may be said to treat others *unjustly*, in this sense, when she violates this principle. For example, a physician who consistently distributes medical service among the ill on the basis of medical need would, *ceteris paribus*, be acting justly in this sense; whereas one who distributes such service without regard to medical needs, but instead with regard to race or religion, would, in this sense, be acting unjustly. This is so because we typically regard medical need as the controlling factor in distributing health care; whereas race and religion appear to be quite irrelevant in such a context.

We may, however, be said to treat others justly when we are respectful of their legal and moral rights, or when we give to them what they rightfully deserve; and we may be said to treat others unjustly when we intrude upon their legal or moral rights, or when we treat them in ways in which they do not deserve to be treated. For example, one acts justly, in this sense, when he keeps an agreement with an individual who has the right to insist upon its being kept; or a judge acts justly, in this sense, when he hands down a well-deserved punishment to a legal offender; whereas a person perpetrates an injustice upon another, in this sense, when he fails to uphold a binding agreement or when he inflicts injury upon an innocent party.

Let us say, then, that the just person is one who is disposed toward treating others justly in both of the above senses. That is, she tends to be consistent in her treatment of others—she does not normally make biased or arbitrary exceptions. But she is also the sort of person who respects individual rights and can usually be counted upon to make good on her obligations to others.

2. Being morally good would also appear to require being *truthful*. By truthful person is meant one who is in the habit of asserting things only if he *believes* them to be true. Thus he is in the habit of asserting things with the intention of *informing* his hearers, and not deceiving them, about the truth. An *untruthful* person, on the other hand, is in the habit of asserting things which she *disbelieves*, and this she does with the intention of deceiving her hearers about the truth. Moreover, the untruthful person may deceive not merely through her spoken word, but also by other means. She may, for example, leave false clues or simply remain silent where such measures are calculated to mislead as to the truth. This is not to suggest that such tactics are never justified; it is rather to say that, when they constitute the rule instead of the exception, the person in question has fallen below that level of truthfulness which we should normally require of a morally good person.

3. Being a morally good person would also seem to demand at least *some* measure of *moral courage*. Indeed, it would appear that a person could not be

just or truthful if he did not have any such measure; for it often takes courage to be honest or to do what is just. By a *morally* courageous person we mean a person who is disposed toward doing what he thinks is morally right even when he believes that his doing so means, or is likely to mean, his suffering some substantial hardship. As Aristotle suggests, it is "the mark of a brave man to face things that are, and seem, terrible for a man, because it is noble to do so and disgraceful not to do so."

And, therefore, we can say, along with Aristotle, that a person who endures hardship just for the sake of some reward—such as fame or fortune—or for the sake of avoiding some punishment—such as public disfavor or legal sanctions—is not truly acting courageously in this sense, for he acts not because it is morally right to do so, but to gain a reward or avoid a punishment.

4. The moral quality of a person is no doubt often revealed through her monetary habits. Indeed, for some individuals, the making of money constitutes an end in itself for which they willfully transgress the bounds of morally permissible conduct—for example, the pimp, the drug dealer, the thief, and the hit man. And some—those whom we characterize as being stingy, miserly, tight—cling to their money with such tenacity that they would sooner allow great iniquities to occur than surrender a dollar.

The morally good person, on the other hand, would appear to be one who had developed *morally respectable* monetary habits. Such a person Aristotle calls a *liberal* person, one who, he states, "will both give and spend the right amounts and on the right objects, alike in small things and in great, and that with pleasure; he will also take the right amounts and from the right sources." Following Aristotle, let us say then that a morally good person must also be, to some degree, a *liberal* person.

5. We should also expect a morally good person to be *benevolent.* By this we mean that she is disposed to do good for others when she is reasonably situated, and to do no harm. And the concern she has for the well-being of others does not arise out of some ulterior motive but rather *for its own sake.* Furthermore, she is disposed toward *feeling* certain ways under certain conditions—for example, feeling sorrow over another's misfortune or taking pleasure in another's good fortune or in helping another.

It is not supposed, however, that in order to be a morally good person one must be disposed toward benefiting others at great sacrifice to oneself; nor is it supposed that such a person must go very far to benefit or feel sympathetic toward those who do not stand in any concrete personal relation such as friendship of kinship. Still, a person who does not go an inch to benefit anyone—unless justice demands it—and who sympathizes with no one is perhaps at most a minimally good person. But one who intentionally harms others, as a matter of course, with pleasure or without regret, cannot normally be regarded as being benevolent. Indeed, such is a mark of a malevolent or morally base person.

6. So, too, would we expect to find *trustworthiness* in a morally good person. That is, we should expect such a person to be in a habit of keeping the confidences and agreements which he freely accepts or enters upon. Indeed, the person who breaks faith for no good reason is not just being dishonest; he is also being a "traitor" or a "double-crosser."

This, however, is not to suppose that an individual must *never* breach a trust if he is to be a morally good person. There are, undoubtedly, some extenuating circumstances in which breaking a trust would be the morally right thing to do—as when keeping it involves working some greater injustice upon someone than that involved in breaking it. Nor is it to be supposed that trustworthiness is a *sufficient* condition of being morally good. There may be loyalty among thieves, for instance, but we should not, for that reason alone, take their lot to be morally good.

7. A morally good person, I would suggest, is one who is regularly disposed to do her *own* moral thinking—that is, to come to her own decisions about moral issues on the basis of her own moral principles; and then, in turn, to act upon her considered judgment. Kant expressed this fact by saying that the will of a morally good person (that is, a morally good will) is one which is determined "autonomously." Following his usage, let us say then that a morally good person is a person who possesses *moral autonomy.*

Being such a person is undoubtedly no easy matter, for moral decisions are frequently difficult ones to make. For instance, in cases of conflict between one's moral principles, one must weigh one principle against another and then "strike a

moral balance"—as, for instance, in a case where keeping a promise involves inflicting harm upon another; or when, in determining the value of the consequences of an act, one must balance the good consequences against the bad ones. And such determinations are clearly no mere matter of logical deduction. All that one can reasonably be expected to do in such cases is to try one's level best. But it is a mark of a morally autonomous person, and thus of a morally good person, that he actually makes such an earnest effort.

Keeping the foregoing criteria of a morally good person in mind, let us now turn to an analysis of lawyers.

THE PURE LEGAL ADVOCATE CONCEPT

Following one traditional usage, we can say that the concept of a lawyer is a *functional* concept—that is, it may be defined in terms of the function or role which a lawyer *qua* lawyer is supposed to perform, in an analogous manner in which a watchdog may be defined in terms of its function of guarding property, or in which a carpenter's hammer may be defined in terms of its function of driving in nails. Hence, just as a good (effective) watchdog may be defined as a dog which performs well the function of guarding property so too may a good (effective) lawyer be defined as a person who performs well the function or role of a lawyer. What, then, we may ask, is the function or role of a lawyer?

One sense of *lawyer* is that in which the role of a lawyer is restricted to that of the client's legal advocate, and in which a good lawyer is thus conceived as being *simply* an effective legal advocate. This sense, which we shall hereafter call the *pure legal advocate concept*, is exemplified in the classic statement made by Lord Brougham when he was defending Queen Caroline against George IV in their divorce case before the House of Lords. He states:

> An advocate, in discharge of his duty, knows but one person in all the world, and that person is his client. To save that client by all means and expedients, and at all hazards and costs to other persons, and, amongst them, to himself, is his first and only duty.

The pure legal advocate concept is also more recently suggested by Canon 15 of the ABA *Canons of Professional Ethics*, which states that

> the lawyer owes "entire devotion to the interest of the client, warm zeal in the maintenance and defense of his rights and the exertion of his utmost learning and ability," to the end that nothing be taken or be withheld from him, save by the rules of law, legally applied.

. . . Given the pure legal advocate concept, it is easy for one to conclude that the necessary and sufficient mark of a good lawyer is her tendency to win cases by all legal means. For, as was said, this concept supposes that a good lawyer is simply an effective legal advocate; and it is easy to suppose that the necessary and sufficient mark of an effective legal advocate is her tendency to legally win cases. A good lawyer hence emerges as a legal technician skillful in manipulating legal rules for the advancement of her clients' legal interests; in this sense, the good lawyer is no different than a skillful chess player able to manipulate the rules of chess to win *his* game.

Furthermore, given this concept, a lawyer may, and indeed is required, to do certain kinds of things on behalf of his client which would ordinarily be regarded as being morally objectionable. In such instances all that matters, so far as lawyering is concerned, is that such acts are legal means of advancing the client's legal interests. For instance, a defense attorney in a rape case may cross-examine the prosecutrix, whom he knows to be telling the truth, about her chastity for purposes of casting doubt upon her truthful testimony. Or, he may permit his client to take the stand knowing full well that the client will perjure himself. Or, a lawyer in a civil case may invoke a legal technically (for example, the statute of limitations) on behalf of his client in order to defeat a just cause against him. Or, a corporate lawyer on a continuing retainer may represent a client who seeks to keep a factory in operation which creates a public health hazard by emitting harmful pollutants into the air.

However, some who countenance the pure legal advocate concept—namely, those sometimes referred to as rule utilitarians—hold that such immoralities as the above-mentioned ones are the necessary evils of

maintaining an adversary system which itself does the greatest good. The working assumption here is that the adversarial form of legal administration, wherein two zealous advocates are pitted against each other before an impartial judge, constitutes the best-known way of maximizing truth and justice; and that, furthermore, this system works best when lawyers disregard their personal moral convictions and hereby restrict their professional activities to the zealous legal representation of their clients.

If the rule utilitarian is correct, then lawyering, so conceived, can be said to be a morally justified function, notwithstanding that, on that view, a lawyer may be required to engage in conduct which, by common standards, is morally objectionable. Thus, when seen in this light, the lawyer emerges as a promoter of the highly prized ends of justice and truth, and as an individual who, because of her service to society, is worthy of praise and admiration. Indeed, she begins to seem like a morally good person.

Nevertheless, I want to suggest that the appearance is deceiving, that, on the contrary, the lawyer, so conceived, will inevitably fall short of our marks of a morally good person. Moreover, I want to suggest that, as a result of such shortcomings, there is substantial disutility in the pure legal advocate concept of lawyering which its utilitarian exponents rarely take into account in their utilitarian justification of it.

Let me emphasize that I am supposing, along with Aristotle, that is takes exercise and training to cultivate the character traits of a morally good person: one is not simply born with them. My claim is, accordingly, that the legal function as construed under the pure legal advocate concept, with its emphasis on suppression of the individual lawyer's personal moral convictions, does not allow for the cultivation of these traits and is, in fact, quite conducive to their corresponding vices.

Furthermore, I am supposing that a lawyer cannot easily detach his professional life from his private life and thereby cannot easily be one sort of person with one set of values in the one life, and a quite different sort with quite different values in the other life.

The latter supposition is justified by the substantial amount of empirical evidence that now exists correlating the personality traits of individuals with their specific vocations. One ambiguity, however, is whether individual vocations influence personality traits, or personality traits influence choice of vocation, or some combination of both; for any one of these hypotheses would explain the correlation.

Some studies have supported the hypothesis that personality traits influence choice of profession—that is, that people with certain personalities are attracted to certain professions in order to satisfy their individual needs. But even *if* this hypothesis is true, and the other above-mentioned hypotheses are false, it is clear that the kind of person found in a profession will remain a function of the way the profession itself is conceived. Specifically, on the hypothesis in question: we would expect the personalities of those choosing careers in law to depend upon their conception of a lawyer. But, if I am correct, then legal practice as construed on the pure legal advocate model could seem attractive only to those individuals contemplating a career in law who would feel comfortable in a professional climate which discourages, rather than promotes, the personality traits of a morally good person as here understood.

MORAL SHORTCOMINGS OF THE PURE LEGAL ADVOCATE

1. It appears that a lawyer, on the pure legal advocate concept, will inevitably fall short of being a *just* person. For although she does not violate the principle of treating relevantly similar cases similarly when she gives special preference to her client—inasmuch as being a client would appear to be a relevant dissimilarity for the purposes of an adversary system—she does, indeed, work injustices through the violation of the moral *rights* of individuals. For on this concept the lawyer's fundamental professional obligation is to do whatever she can, within legal limits, to advance the legal interests of her clients. But from this basic obligation there derives a more specific one which, contra Kant, may be expressed thus: "Whenever legally possible treat others not as ends but as means toward winning your case."

For example, the criminal defense lawyer is thereby authorized to knowingly destroy the testimony of an innocent rape victim in order to get an acquittal for his client; and a civil lawyer is authorized to knowingly deprive another of what he rightfully deserves by invoking the statute of limitations

for the purpose of furthering his client's interests. But we shall concur with Kant in maintaining that lawyers, like anyone else, have a duty to treat others with the respect which they, as persons, have a right to insist upon.

2. Nor will the pure legal advocate meet the mark of *truthfulness.* For, from her cardinal obligation there derives the secondary obligation of being untruthful where doing so can legally contribute toward winning the case. An example of a lawyer who complies with this obligation is one who remains silent when she knows that her client has, under oath, lied to the court. The lawyer, by wittingly saying nothing, engages in deceptive behavior—she contributes to the court's being deceived as to the truth—and is on that count *herself* guilty of being untruthful. Indeed, scrupulous adherence to *this* obligation could hardly support anything but an untruthful habit.

3. Nor does the concept in question support *moral courage.* For, according to it, the personal moral convictions of a lawyer are irrelevant to his function and should not serve as reasons for zealous representation of clients, or for any sacrifices—of time, money, reputation, and the like—which he may make on their behalf. Indeed, if he is to do his job well, then he must get into the habit of *not* being influenced by his moral outlook. Rather, any sacrifice he may make should be for the sake of obtaining a legal victory, be it a moral one or not. It is plausible to suppose, however, that where morality takes a back seat, ulterior motives, such as the self-aggrandizement obtained through winning, will server as the primary motivation.

4. Nor does the pure legal advocate concept support *liberality*; for the pure legal advocate, through her unconcern with the moral character of her clients and the purposes for which they hire her, gets into the habit of taking money from dishonorable individuals for unsavory purposes. She thus emerges as a professional who can be hired, for a good sum, to do the dirty work of a villain or a scoundrel. Indeed, she then begins to sound more like a hired assassin than like the liberal person whom Aristotle had in view. The high-priced corporate lawyer who wittingly helps her corporate client to market a dangerous product provides us with one example of such a lawyer, and the high-priced criminal lawyer who specializes in defending mass murderers is another.

5. Furthermore, the pure legal advocate concept does not appear to satisfy the minimum condition of *benevolence*—that is, the nonmalevolence expected of a morally good person. For, from his primary obligation, there derives the secondary obligation to employ even such means to forward a client's interests as are injurious to others, so long, of course, as they are legal. But this also means that a lawyer must learn to put off sympathetic feelings which a benevolent person would normally have. In particular, he must get used to working injury upon others without having any strong feelings of guilt, sorrow, or regret. For, to be sure, such feelings could only serve to interfere with the execution of his basic obligation to his client. The result is thus a callous attitude in his dealings with others.

6. Prime facie, it appears that the morally desirable character trait of *trustworthiness* receives strong support from the pure legal advocate concept. For, indeed, it appears that a lawyer cannot put on the most effective representation of her client's interests unless she is also prepared to hold in confidence the secrets entrusted to her by her client. A problem with this view, however, arises in the case in which there is a conflict between a lawyer's obligation to keep her client's confidence and some other moral obligation—for instance, that of not harming innocent persons. In such a case, the restricted lawyer is required to keep her client's confidence so long as it is legally possible and in her client's best interest to do so. Her considered judgment as to what is, under the circumstances, morally best is then quite irrelevant. But it is a mark of a morally good person to choose what she thinks is, all things considered, the morally right thing to do in such a situation. Hence, whereas a morally good person sees his obligation to keep confidences as one among several moral principles which may at times override one another, the pure legal advocate sees her professional obligation to keep her clients' confidences as binding upon her quite independently of the moral propriety of doing so in any particular case.

In any event, even *if* it is admitted that the pure legal advocate concept reinforces trustworthiness which, *in itself*, is a morally good trait, this still does not show that the good lawyer, on this conception,

can be a morally good person. For, as we have seen, there are further requisites of a morally good life.

7. I have suggested that an important quality of a morally good person is that he has *moral autonomy.* However, the pure legal advocate concept offers no stimulus to the cultivation of this trait. For, as we have seen, the pure legal advocate inhabits a world in which his moral judgment is quite beside the point. If morality is relevant, it is so at the level of the judge or the legislator, but it is quite outside the purview of the lawyer's function. The lawyer must know the law and must know that he owes his undivided allegiance to his client. Given the latter he can easily accommodate himself to the requirements of the law. His decisions are, in effect, made *for him* by the system he serves. He is more like a cog in a machine and less like a person. But the moral world is inhabited by *persons*—that is, individuals who autonomously confront their moral responsibilities; so that, for a lawyer who has grown comfortable with passing the buck of moral responsibility, there is little hope of his aspiring to the morally good life.

THE MORAL AGENT CONCEPT

If I am correct, then it appears that the pure legal advocate who scrupulously adheres to her restricted role, far from being a morally good person, will be given ample opportunity for becoming—if she is not already—quite the opposite. For she will thereby be placed in a professional climate conducive to her being unjust instead of just; untruthful instead of truthful; unmotivated by a moral outlook instead of morally courageous; illiberal instead of liberal; callous instead of benevolent; morally irresponsible instead of morally autonomous. In short, she will fall well below the minimum standards of a morally good person.

But if all this is right, then there will, it seems, be a good deal of *disutility* in the pure legal advocate concept which, indeed, any utilitarian exponent of it ought to consider in computing its overall balance of utility. For it appears that such personality traits as those mentioned above, when associated with our concept of a lawyer, can serve only to bring disrespect upon the legal profession and, by association, upon the legal system as a whole. And this low regard may well lead to a commonplace view of the adversary system as a haven for the guilty and the wicked and as something of which the innocent and the morally good ought to steer clear. It can very well serve to discourage persons of strong moral character from entering the legal profession. It is also quite plausible that pure legal advocates who, by virtue of their knowledge of, and relation to, the law, are uniquely situated to contribute to needed changes in unjust laws, will not concern themselves with such moral reformation. Moreover, add to these the disutility involved in the unsavory acts performed by pure legal advocates on behalf of their clients in the normal course of discharging their professional obligations—the injuries thereby done to individual litigants as well as to others—and there is at least a strong prima facie case for abandoning the adversary system entirely in favor of a different model (an inquisitorial model for instance) or for adopting a concept of a lawyer in an adversary system which avoids these disutilities.

Fortunately, there is a further concept of a lawyer which, while not abandoning the adversarial approach, serves to avoid much of the disutility mentioned above.

This farther sense, hereafter called the *moral agent concept*, is exemplified, for example, in the remarks on advocacy made by Lord Chief Justice Cockburn, in the presence of Lord Brougham, at a dinner given in honor of M. Berryer on November 8, 1864. He stated:

> My noble and learned friend, Lord Brougham, whose words are the words of wisdom, said that an advocate should be fearless in carrying out the interests of his client: but I couple that with this qualification and this restriction—that the arms which he wields are to be the arms of the warrior and not of the assassin. It is his duty to strive to accomplish the interest of his clients per fas, but not per nefas; it is his duty, to the utmost of his power, to seek to reconcile the interests he is bound to maintain, and the duty it is incumbent upon him to discharge, with the eternal and immutable interests of truth and justice.

The moral agent concept was also expressed, more recently, by John Noonan when he remarked that

> a lawyer should not impose his conscience on his client; neither can he accept his client's decision and remain entirely free from all moral responsibility,

> subject only to the restraints of the criminal law. The framework of the adversary system provides only the first set of guidelines for a lawyer's conduct. He is also a human being and cannot submerge his humanity by playing a technician's role.

And this concept is suggested elsewhere by Richard Wasserstrom, Jeremy Bentham, and *The Report of the Joint Conference on Professional Responsibility.*

Given the moral agent concept, we may no longer say that the good lawyer is simply the effective legal advocate; he is, rather, one who is effective in morally as well as legally advocating his client's cause. Hence, one cannot infer from this concept that the good lawyer is one who tends to win his cases. For, on this concept, he is not merely a good legal technician; he is also one who conducts himself in the manner of a morally good person—that is, as a person with morally desirable character traits.

It is evident, however, that a lawyer cannot so conduct herself unless she also subscribes to the moral principles to which a morally good person would subscribe were she to participate in an adversarial process. If our analysis of a morally good person is supposed, then such principles would need to be ones supportive of the personality traits set forth in that analysis. To wit, from these character traits we may derive a corresponding set of moral principles which are adjusted to an adversarial context. I suggest the following formulations, although other similar formulations are possible.

- Treat others as ends in themselves and not as mere means to winning cases. (Principle of Individual Justice)
- Treat clients and other professional relations who are relevantly similar in a similar fashion. (Principle of Distributive Justice)
- Do not deliberately engage in behavior apt to deceive the court as to the truth. (Principle of Truthfulness)
- Be willing, if necessary, to make reasonable personal sacrifices—of time, money, popularity, and so on—for what you justifiably believe to be a morally good cause. (Principle of Moral Courage)
- Do not give money to, or accept money from, clients for wrongful purposes or in wrongful amounts. (Principle of Liberality)
- Avoid harming others in the process of representing your client. (Principle of Nonmalevolence)
- Be loyal to your client, and do not betray his confidences. (Principles of Trustworthiness)
- Make your *own* moral decisions to the best of your ability and act consistently upon them. (Principle of Moral Autonomy)

We can say that above principles, or ones like them, at least in part *constitute* or *define* the moral agent concept of a lawyer; for they are principles to which a lawyer's conduct must to some extent conform if he is to function not simply as a legal advocate but also as a morally good person.

I am *not* suggesting that these principles are unconditional ones. Indeed, to say so would be unrealistic since they will inevitably come into conflict with each other when applied to specific contexts, thereby making it impossible for the lawyer to satisfy all principles at once. (That is, in order to be truthful, a lawyer may need to betray a client's trust, and conversely.) Rather, what I am suggesting is that such principles impose upon a lawyer *conditional*—or prima facie—obligations which, in cases of conflict, must be weighed, one against the other, by the lawyer in question in the context in question.

Let me offer an example which will illustrate the difference between applying, in conflict situations, the above multi-principle model and the pure legal advocate model. In *Lawyers' Ethics in an Adversary System*, Monroe Freedman cites the following:

> In a recent case in Lake Pleasant, New York, a defendant in a murder case told his lawyers about two other people he had killed and where their bodies had been hidden. The lawyers went there, observed the bodies, and took photographs of them. They did not, however, inform the authorities about the bodies until several months later, when their client had confessed to those crimes. In addition to withholding the information from police and prosecutors, one of the attorneys denied information to one of the victims' parents, who came to him in the course of seeking his missing daughter.

According to Freedman, the lawyers in the above cited case were simply discharging their *unconditional* professional obligation to represent their clients' legal interests. However, if the moral agent concept is supposed, then it is clear that the above lawyers could have revealed where the

bodies were buried. Admittedly, according to the Principle of Trustworthiness, a lawyer has a (prima facie) obligation to keep his client's confidences. But he also has further (prima facie) obligation such as those of Truthfulness, Individual Justice, and Nonmalevolence. I think that a plausible case can be made that the latter principles were sacrificed to some extent by the lawyers in the Lake Pleasant case at least insofar as their treatment of the relatives of the deceased was concerned. And it is plausible, I believe, to argue that the moral weight of the latter principles, taken collectively, outweighed that of the Principle of Trustworthiness taken by itself in the situation in question. This need not have been what the lawyers in the cited case should have finally decided to be the correct balancing of principles. The point I want to make is rather that the lawyers *did have* in the first place, on the conception in question, the *moral autonomy* (as legitimized by our Principle of Moral Autonomy) to make such a judgment on the matter. It is just such moral autonomy—with its weighing of competing moral principles, one against the other—that the pure legal advocate concept disallows.

Of course, if lawyers are allowed such autonomy, there arises the difficulty of providing *criteria* for arbitrating between conflicting principles. This difficulty, we have seen, does not arise on the pure legal advocate model since the pure legal advocate is, in effect, insulated from making moral trade-offs by her unconditional allegiance to her client's legal interests.

One normative view regarding how a lawyer, in accordance with the moral agent concept, might go about solving moral dilemmas takes the form of a "pure" utilitarianism. According to such an ethic, all eight of our principles are to be understood as receiving their ultimate justification from the principle of utility. Hence, in case of conflict the final court of appeal will be the principle of utility itself.

I do not think, however, that such a basis for solving lawyers' moral dilemmas would be adequate. My objection is that which has traditionally been made against utilitarian ethics which are not tempered by justice considerations. Suppose, for example, a criminal lawyer is defending an influential politician accused of rape. Suppose also that the politician admits his guilt to his attorney but nevertheless informs her of his intention to testify under oath (to that which is false) that the defendant first made sexual advances toward him. Now suppose that the politician in question is in the process of bringing about a change in taxation which would mean substantial tax reductions for millions of Americans, and that, furthermore, these efforts would most likely be defeated if the politician in question were convicted of rape. On the pure utilitarian criterion, it would appear that the attorney in question would be committed to allowing the politician to perjure himself notwithstanding the defeat of the true rape claim; for the greatest good would (*ex hypothesi*) be served by allowing the politician to escape the charge of rape through his perjured testimony. But in such a case it would seem unjust (by the Principle of Individual Justice) to sacrifice the well-being of the truthful rape victim for the tax reduction. Indeed, in doing so, the lawyer would arguably be committing a grossly immoral act. But, if so, the principle of utility untempered by some principle(s) of justice—such as the Principle of Individual Justice or the Principle of Distributive Justice—would be an inadequate criterion for settling lawyers' moral dilemmas. . . .

How then is a lawyer, on the moral agent conception, to resolve antinomies arising between those principles? Although I do not see any formula for doing so, this is not to suggest that one resolution is just as respectable as any other. For one thing, there is a difference between the ethical judgment of a lawyer that is *factually enlightened* and one that is not. For example, the judgment of lawyer who allows trustworthiness to override harm in a particular case without adequate knowledge of the nature and extent of the harm is less respectable than the judgment of a lawyer who takes account of such facts.

Still, once the facts are known a decision must be made in their light; and I think it would be intellectually dishonest to suggest that there is some principle(s) from which we may logically deduce our decision. Principles take us just so far, leaving the final verdict in our hands.

It is the lack of clear noncontroversial criteria for resolving moral dilemmas, and the ensuing feeling that ethics is, in the end, a matter of "fiat" or "personal preference," that may make some feel

uncomfortable about giving lawyers moral autonomy. However, it should be kept in mind by those who worry about the "gray" areas of ethics that the making of ethical decisions is already an accepted and unavoidable part of the role of *some* officials in our legal system. For example, given the "open-textured" quality of legal rules and precedents themselves, judges often need to rely upon their own *moral* evaluations in deciding whether a given set of facts falls under a given legal rule or precedent. But if judges can handle their moral problems—and I believe that, in general, they do handle them—then there appears to be less reason to fear that lawyers cannot or will not handle their moral problems.

REVIEW AND DISCUSSION QUESTIONS

1. According to Cohen, what seven characteristics do we associate with a morally good person? Do you agree with his list? In your view, which characteristics are the most important? Are there characteristics that Cohen has overlooked or that could be dropped from his list?
2. What is the pure legal advocate concept, and what do you see as its implications for legal practice? Can the concept be defended from a rule-utilitarian point of view? Explain why or why not.
3. Reviewing each of the traits that characterize a morally good person, Cohen argues that a pure legal advocate will not be a morally good person. For each trait, explain whether you agree or disagree with Cohen. In your opinion, is it possible to be a good lawyer and a good person?
4. Do you agree or disagree with Cohen that there is "a great deal of *disutility* in the pure legal advocate concept"?
5. Cohen identifies eight principles that constitute, at least in part, the moral agent concept of a lawyer. In your view, which of these principles are the most important? Are there other principles you would add to Cohen's list?
6. Cohen says that the obligations imposed by his principles are conditional or prima facie. How does this point apply to Freedman's case of the hidden bodies?
7. How is a lawyer to resolve possible conflicts between Cohen's principles? Can utilitarianism or some other ethical theory provide the answer?
8. How difficult do you think it is to be an effective and successful lawyer and yet remain a morally good person? To what extent will this be affected by the kind of law one practices or the kind of clients one represents?

CHAPTER 2

The Rule of Law

This section focuses on a theme that has long been of interest to philosophers and legal theorists: the idea of the rule of law. In their essays, A. V. Dicey, Lon Fuller, Joseph Raz and H. L. A. Hart, illuminate different aspects of the rule of law, examining its meaning and limits, how it can best be protected, and why maintaining it is important. These general issues are brought to bear on two more specific puzzles: "The Problem of the Grudge Informer" concerns how a newly chosen, democratic government should deal with those who used the authority of a previous Nazi-like regime for personal advantage, while *Hamdi v. Rumsfeld* probes the legality of detaining American citizens as "enemy combatants" and the requirements of habeas corpus in a time of war.

Rule of Law and the English Constitution

A. V. Dicey

In this excerpt from his classic work on English constitutional law, the jurist and legal scholar A. V. Dicey expounds what he takes to be the central features of the rule of law as exemplified in the English constitution: (1) Society is governed by law, not arbitrary power, and no one may be punished except for a breach of law; (2) there is equality before the law in the sense that everyone, regardless of class, is subject to the ordinary law of the land; and (3) the general principles of the constitution, in particular, the rights of individuals, are a result of judicial decisions in particular cases brought before the courts rather than (as in other systems) deductions from general declarations and definitions of rights. The excerpt concludes with a discussion of the writ of habeas corpus, long held to be a cornerstone of liberty and the rule of law in England and the United States.

Two features have at all times since the Norman Conquest characterised the political institutions of England.

The first of these features is the omnipotence or undisputed supremacy throughout the whole country of the central government. This authority of the

state or the nation was during the earlier periods of our history represented by the power of the Crown. The King was the source of law and the maintainer of order. . . . This royal supremacy has now passed into that sovereignty of Parliament which has formed the main subject of the foregoing chapters.

The second of these features, which is closely connected with the first, is the rule or supremacy of law. . . .

This supremacy of the law, or the security given under the English constitution to the rights of individuals looked at from various points of view, forms the subject of this part of this treatise.

When we say that the supremacy or the rule of law is a characteristic of the English constitution, we generally include under one expression at least three distinct though kindred conceptions.

We mean, in the first place, that no man is punishable or can be lawfully made to suffer in body or goods except for a distinct breach of law established in the ordinary legal manner before the ordinary Courts of the land. In this sense the rule of law is contrasted with every system of government based on the exercise by persons in authority of wide, arbitrary, or discretionary powers of constraint.

Modern Englishmen may at first feel some surprise that the "rule of law" (in the sense in which we are now using the term) should be considered as in any way a peculiarity of English institutions, since, at the present day, it may seem to be not so much the property of any one nation as a trait common to every civilised and orderly state. Yet, even if we confine our observation to the existing condition of Europe, we shall soon be convinced that the "rule of law" even in this narrow sense is peculiar to England, or to those countries which, like the United States of America, have inherited English traditions. In almost every continental community the executive exercises far wider discretionary authority in the matter of arrest, of temporary imprisonment, of expulsion from its territory, and the like, than is either legally claimed or in fact exerted by the government in England; and a study of European politics now and again reminds English readers that wherever there is discretion there is room for arbitrariness, and that in a republic no less than under a monarchy discretionary authority on the part of the government must mean insecurity for legal freedom on the part of its subjects.

If, however, we confined our observation to the Europe of the twentieth century, we might well say that in most European countries the rule of law is now nearly as well established as in England, and that private individuals at any rate who do not meddle in politics have little to fear, as long as they keep the law, either from the Government or from any one else. . . .

We mean in the second place, when we speak of the "rule of law" as a characteristic of our country, not only that with us no man is above the law, but (what is a different thing) that here every man, whatever be his rank or condition, is subject to the ordinary law of the realm and amenable to the jurisdiction of the ordinary tribunals.

In England the idea of legal equality, or of the universal subjection of all classes to one law administered by the ordinary Courts, has been pushed to its utmost limit. With us every official, from the Prime Minister down to a constable or a collector of taxes, is under the same responsibility for every act done without legal justification as any other citizen. The Reports abound with cases in which officials have been brought before the Courts, and made, in their personal capacity, liable to punishment, or to the payment of damages, for acts done in their official character but in excess of their lawful authority. . . .

There remains yet a third and a different sense in which the "rule of law" or the predominance of the legal spirit may be described as a special attribute of English institutions. We may say that the constitution is pervaded by the rule of law on the ground that the general principles of the constitution (as for example the right to personal liberty, or the right of public meeting) are with us the result of judicial decisions determining the rights of private persons in particular cases brought before the Courts; whereas under many foreign constitutions the security (such as it is) given to the rights of individuals results, or appears to result, from the general principles of the constitution. . . .

Our constitution, in short, is a judge-made constitution, and it bears on its face all the features, good and bad, of judge-made law. Hence flow noteworthy distinctions between the constitution of England and the constitutions of most foreign countries.

There is in the English constitution an absence of those declarations or definitions of rights so dear to foreign constitutionalists. Such principles, moreover, as you can discover in the English constitution are, like all maxims established by judicial legislation, mere generalisations drawn either from the decisions or dicta of judges, or from statutes which, being passed to meet special grievances, bear a close resemblance to judicial decisions. . . . To put what is really the same thing in a somewhat different shape, the relation of the rights of individuals to the principles of the constitution is not quite the same in countries like Belgium, where the constitution is the result of a legislative act, as it is in England, where the constitution itself is based upon legal decisions. In Belgium, which may be taken as a type of countries possessing a constitution formed by a deliberate act of legislation, you may say with truth that the rights of individuals to personal liberty flow from or are secured by the constitution. In England the right to individual liberty is part of the constitution, because it is secured by the decisions of the Courts, extended or confirmed as they are by the *Habeas Corpus* Acts. . . . [I]n Belgium individual rights are deductions drawn from the principles of the constitution, whilst in England the so-called principles of the constitution are inductions or generalisations based upon particular decisions pronounced by the Courts as to the rights of given individuals.

This is of course a merely formal difference. Liberty is as well secured in Belgium as in England, and as long as this is so it matters nothing whether we say that individuals are free from all risk of arbitrary arrest, because liberty of person is guaranteed by the constitution, or that the right to personal freedom, or in other words to protection from arbitrary arrest, forms part of the constitution because it is secured by the ordinary law of the land. But though this merely formal distinction is in itself of no moment, provided always that the rights of individuals are really secure, the question whether the right to personal freedom or the right to freedom of worship is likely to be secure does depend a good deal upon the answer to the inquiry whether the persons who consciously or unconsciously build up the constitution of their country begin with definitions or declarations of right, or with the contrivance of remedies by which rights may be enforced or secured.

Now, most foreign constitution-makers have begun with declarations of rights. . . . But any knowledge of history suffices to show that foreign constitutionalists have, while occupied in defining rights, given insufficient attention to the absolute necessity for the provision of adequate remedies by which the rights they proclaimed might be enforced. . . . On the other hand, there runs through the English constitution that inseparable connection between the means of enforcing a right and the right to be enforced which is the strength of judicial legislation. . . . The *Habeas Corpus* Acts declare no principle and define no rights, but they are for practical purposes worth a hundred constitutional articles guaranteeing individual liberty. Nor let it be supposed that this connection between rights and remedies which depends upon the spirit of law pervading English institutions is inconsistent with the existence of a written constitution, or even with the existence of constitutional declarations of rights. The Constitution of the United States and the constitutions of the separate States are embodied in written or printed documents, and contain declarations of rights. But the statesmen of America have shown unrivalled skill in providing means for giving legal security to the rights declared by American constitutions. The rule of law is as marked a feature of the United States as of England.

The fact, again, that in many foreign countries the rights of individuals, *e.g.* to personal freedom, depend upon the constitution, whilst in England the law of the constitution is little else than a generalisation of the rights which the Courts secure to individuals, has this important result. The general rights guaranteed by the constitution may be, and in foreign countries constantly are, suspended. They are something extraneous to and independent of the ordinary course of the law. The declaration of the Belgian constitution, that individual liberty is "guaranteed," betrays a way of looking at the rights of individuals very different from the way in which such rights are regarded by English lawyers. . . . Where, on the other hand, the right to individual freedom is part of the constitution because it is inherent in the ordinary law of the land, the right is one which can hardly be destroyed without a thorough revolution in the institutions and manners of the nation. The so-called "suspension of

the *Habeas Corpus* Act" bears, it is true, a certain similarity to what is called in foreign countries "suspending the constitutional guarantees." But, after all, a statute suspending the *Habeas Corpus* Act falls very far short of what its popular name seems to imply; and though a serious measure enough, is not, in reality, more than a suspension of one particular remedy for the protection of personal freedom. The *Habeas Corpus* Act may be suspended and yet Englishmen may enjoy almost all the rights of citizens. . . .

That "rule of law," then, which forms a fundamental principle of the constitution, has three meanings, or may be regarded from three different points of view.

It means, in the first place, the absolute supremacy or predominance of regular law as opposed to the influence of arbitrary power, and excludes the existence of arbitrariness, of prerogative, or even of wide discretionary authority on the part of the government. Englishmen are ruled by the law, and by the law alone; a man may with us be punished for a breach of law, but he can be punished for nothing else.

It means, again, equality before the law, or the equal subjection of all classes to the ordinary law of the land administered by the ordinary Law Courts; the "rule of law" in this sense excludes the idea of any exemption of officials or others from the duty of obedience to the law which governs other citizens or from the jurisdiction of the ordinary tribunals; there can be with us nothing really corresponding to the "administrative law" (*droit administratif*) or the "administrative tribunals" (*tribunaux administratifs*) of France. The notion which lies at the bottom of the "administrative law" known to foreign countries is, that affairs or disputes in which the government or its servants are concerned are beyond the sphere of the civil Courts and must be dealt with by special and more or less official bodies. This idea is utterly unknown to the law of England, and indeed is fundamentally inconsistent with our traditions, and customs.

The "rule of law," lastly, may be used as a formula for expressing the fact that with us the law of the constitution, the rules which in foreign countries naturally form part of a constitutional code, are not the source but the consequence of the rights of individuals, as defined and enforced by the Courts; that, in short, the principles of private law have with us been by the action of the Courts and Parliament so extended as to determine the position of the Crown and of its servants; thus the constitution is the result of the ordinary law of the land.

General propositions, however, as to the nature of the rule of law carry us but a very little way. If we want to understand what that principle in all its different aspects and developments really means, we must try to trace its influence throughout some of the main provisions of the constitution. . . .

WRIT OF HABEAS CORPUS

The right of a person who has been wrongfully imprisoned on regaining his freedom to put his oppressor on trial as a criminal, or by means of an action to obtain pecuniary compensation for the wrong which he has endured, affords a most insufficient security for personal freedom. If *X* keeps *A* in confinement, it profits *A* little to know that if he could recover his freedom, which he cannot, he could punish and fine *X*. What *A* wants is to recover his liberty. Till this is done he cannot hope to punish the foe who has deprived him of it. . . . Liberty is not secure unless the law, in addition to punishing every kind of interference with a man's lawful freedom, provides adequate security that every one who without legal justification is placed in confinement shall be able to get free. This security is provided by the celebrated writ of *habeas corpus* and the *Habeas Corpus* Acts.

It is not within the scope of these lectures to give a history of the writ of *habeas corpus* or to provide the details of the legislation with regard to it. . . . My object is solely to explain generally the mode in which the law of England secures the right to personal freedom. . . .

Nature of Writ

Legal documents constantly give the best explanation and illustration of legal principles. We shall do well therefore to examine with care the following copy of a writ of *habeas corpus*:

> *Victoria, by the Grace of God, of the United Kingdom of Great Britain and Ireland Queen, Defender of the Faith,*
>
> *To J. K., Keeper of our Gaol of Jersey, in the Island of Jersey, and to J. C. Viscount of said Island, greeting.*

We command you that you have the body of C. C. W. detained in our prison under your custody, as it is said, together with the day and cause of his being taken and detained, by whatsoever name he may be called or known, in our Court before us, at Westminster, on the 18th day of January next, to undergo and receive all and singular such matters and things which our said Court shall them and there consider of him in this behalf; and have there then this Writ. Witness THOMAS *Lord* DENMAN, *at Westminster, the 23rd day of December in the 8th year of our reign.*

By the Court,
Robinson.
R.M.R.
At the instance of C. C. W.
W. A. L., 7 Gray's Inn Square, London,
Attorney for the said C. C. W.

The character of the document is patent on its face. It is an order issued, in the particular instance, by the Court of Queen's Bench, calling upon a person by whom a prisoner is alleged to be kept in confinement to bring such prisoner—to "have his body," whence the name *habeas corpus*—before the Court to let the Court know on what ground the prisoner is confined, and thus to give the Court the opportunity of dealing with the prisoner as the law may require. The essence of the whole transaction is that the Court can by the writ of *habeas corpus* cause any person who is imprisoned to be actually brought before the Court and obtain knowledge of the reason why he is imprisoned; and then having him before the Court, either then and there set him free or else see that he is dealt with in whatever way the law requires, as, for example, brought speedily to trial.

The writ can be issued on the application either of the prisoner himself or of any person on his behalf, or (supposing the prisoner cannot act) then on the application of any person who believes him to be unlawfully imprisoned. It is issued by the High Court, or during vacation by any judge thereof; and the Court or a judge should and will always cause it to be issued on being satisfied by affidavit that there is reason to suppose a prisoner to be wrongfully deprived of his liberty. You cannot say with strictness that the writ is issued "as a matter of course," for some ground must be shown for supposing that a case of illegal imprisonment exists. But the writ is granted "as a matter of right,"—that is to say, the Court will always issue it if *prima facie* ground is shown for supposing that the person on whose behalf it is asked for is unlawfully deprived of his liberty. The writ or order of the Court can be addressed to any person whatever, be he an official or a private individual, who has, or is supposed to have, another in his custody. Any disobedience to the writ exposes the offender to summary punishment for contempt of Court, and also in many cases to heavy penalties recoverable by the party aggrieved. . . .

REVIEW AND DISCUSSION QUESTIONS

1. What is the first way in which the rule of law is characteristic of the English constitution? What does it mean to say that regular law, not arbitrary power, is supreme? Is Dicey right to assert that rule of law in this sense is a distinctive feature of England and those countries like the United States that have inherited English legal and political traditions? To what extent does the principle that none can be punished except for a breach of law protect citizens against despotic government?
2. What does it mean to say that there is equality before the law? Can that equality be squared with the inequalities of wealth and power that in fact exist among citizens? Can there be equality before the law in an unjust society? Can there be equality before the law in a society governed by a tyrannical government?
3. In Dicey's view, in what important way does the English constitution differ from most foreign constitutions? In your view, which type of constitution does a better job of protecting individual rights or better fulfills the idea of rule of law, or is there no important difference between the constitutions in this regard?
4. Dicey believes that the rule of law is as secure in the United States as it is in England, yet America has a written constitution. How can this be squared with his criticism of foreign constitutions with their abstract declarations of rights?
5. What exactly is the writ of *habeas corpus*? What does it do, and why is it important? Would the rule of law be possible without the writ?

Eight Ways to Fail to Make Law

Lon Fuller

The task of making a law may seem simple, with the only real problem being how to decide which laws to make. But in this tale, Lon Fuller illustrates that making law at all is often a difficult task that may or may not be effectively accomplished. Implicit in this story is another, deeper issue about the nature of law and the ideal of the rule of law. Lon Fuller was professor of law at Harvard Law School.

> [A] law which a man cannot obey, nor act according to it, is void and no law: and it is impossible to obey contradictions, or act according to them.
>
> —Vaughan, C. J. in *Thomas v. Sorrell,* 1677

> It is desired that our learned lawyers would answer these ensuing queries . . . whether ever the Commonwealth, when they chose the Parliament, gave them a lawless unlimited power, and at their pleasure to walk contrary to their own laws and ordinances before they have repealed them?
>
> —Lilburne, *England's Birth-Right Justified,* 1645

This chapter will begin with a fairly lengthy allegory. It concerns the unhappy reign of a monarch who bore the convenient, but not very imaginative and not even very regal sounding name of Rex.

EIGHT WAYS TO FAIL TO MAKE LAW

Rex came to the throne filled with the zeal of a reformer. He considered that the greatest failure of his predecessors had been in the field of law. For generations the legal system had known nothing like a basic reform. Procedures of trial were cumbersome, the rules of law spoke in the archaic tongue of another age, justice was expensive, the judges were slovenly and sometimes corrupt. Rex was resolved to remedy all this and to make his name in history as a great lawgiver. It was his unhappy fate to fail in this ambition. Indeed, he failed spectacularly, since not only did he not succeed in introducing the needed reforms, but he never even succeeded in creating any law at all, good or bad.

His first official act was, however, dramatic and propitious. Since he needed a clean slate on which to write, he announced to his subjects the immediate repeal of all existing law, of whatever kind. He then set about drafting a new code. Unfortunately, trained as a lonely prince, his education had been very defective. In particular he found himself incapable of making even the simplest generalizations. Though not lacking in confidence when it came to deciding specific controversies, the effort to give articulate reasons for any conclusion strained his capacities to the breaking point.

Becoming aware of his limitations, Rex gave up the project of a code and announced to his subjects that henceforth he would act as a judge in any disputes that might arise among them. In this way under the stimulus of a variety of cases he hoped that his latent powers of generalization might develop and, proceeding case by case, he would gradually work out a system of rules that could be incorporated in a code. Unfortunately the defects in his education were more deep-seated than he had supposed. The venture failed completely. After he had handed down literally hundreds of decisions neither he nor his subjects could detect in those decisions any pattern whatsoever. Such tentatives toward generalization as were to be found in his opinions only compounded the confusion, for they gave false leads to his subjects and threw his own meager powers of judgment off balance in the decision of later cases.

After this fiasco Rex realized it was necessary to take a fresh start. His first move was to subscribe to a course of lessons in generalization. With his intellectual powers thus fortified, he resumed the project of a code and, after many hours of solitary labor, succeeded in preparing a fairly lengthy document. He was still not confident, however, that he had fully overcome his previous defects. Accordingly, he announced to his

subjects that he had written out a code and would henceforth be governed by it in deciding cases, but that for an indefinite future the contents of the code would remain an official state secret, known only to him and his scrivener. To Rex's surprise this sensible plan was deeply resented by his subjects. They declared it was very unpleasant to have one's case decided by rules when there was no way of knowing what those rules were.

Stunned by this rejection Rex undertook an earnest inventory of his personal strengths and weaknesses. He decided that life had taught him one clear lesson, namely, that it is easier to decide things with the aid of hindsight than it is to attempt to foresee and control the future. Not only did hindsight make it easier to decide cases, but—and this was of supreme importance to Rex—it made it easier to give reasons. Deciding to capitalize on this insight, Rex hit on the following plan. At the beginning of each calendar year he would decide all the controversies that had arisen among his subjects during the preceding year. He would accompany his decisions with a full statement of reasons. Naturally, the reasons thus given would be understood as not controlling decisions in future years, for that would be to defeat the whole purpose of the new arrangement, which was to gain the advantages of hindsight. Rex confidently announced the new plan to his subjects, observing that he was going to publish the full text of his judgments with the rules applied by him, thus meeting the chief objection to the old plan. Rex's subjects received this announcement in silence, then quietly explained through their leaders that when they said they needed to know the rules, they meant they needed to know them *in advance* so they could act on them. Rex muttered something to the effect that they might have made that point a little clearer, but said he would see what could be done.

Rex now realized that there was no escape from a published code declaring the rules to be applied in future disputes. Continuing his lessons in generalization, Rex worked diligently on a revised code, and finally announced that it would shortly be published. This announcement was received with universal gratification. The dismay of Rex's subjects was all the more intense, therefore, when his code became available and it was discovered that it was truly a masterpiece of obscurity. Legal experts who studied it declared that there was not a single sentence in it that could be understood either by an ordinary citizen or by a trained lawyer. Indignation became general and soon a picket appeared before the royal palace carrying a sign that read, "How can anybody follow a rule that nobody can understand?"

The code was quickly withdrawn. Recognizing for the first time that he needed assistance, Rex put a staff of experts to work on a revision. He instructed them to leave the substance untouched, but to clarify the expression throughout. The resulting code was a model of clarity, but as it was studied it became apparent that its new clarity had merely brought to light that it was honeycombed with contradictions. It was reliably reported that there was not a single provision in the code that was not nullified by another provision inconsistent with it. A picket again appeared before the royal residence carrying a sign that read, "This time the king made himself clear—in both directions."

Once again the code was withdrawn for revision. By now, however, Rex had lost his patience with his subjects and the negative attitude they seemed to adopt toward everything he tried to do for them. He decided to teach them a lesson and put an end to their carping. He instructed his experts to purge the code of contradictions, but at the same time to stiffen drastically every requirement contained in it and to add a long list of new crimes. Thus, where before the citizen summoned to the throne was given ten days in which to report, in the revision the time was cut to ten seconds. It was made a crime, punishable by ten years' imprisonment, to cough, sneeze, hiccough, faint or fall down in the presence of the king. It was made treason not to understand, believe in, and correctly profess the doctrine of evolutionary, democratic redemption.

When the new code was published a near revolution resulted. Leading citizens declared their intention to flout its provisions. Someone discovered in an ancient author a passage that seemed apt: "To command what cannot be done is not to make law; it is to unmake law, for a command that cannot be obeyed serves no end but confusion, fear and chaos." Soon this passage was being quoted in a hundred petitions to the king.

The code was again withdrawn and a staff of experts charged with the task of revision. Rex's instructions to the experts were that whenever they encountered a rule requiring an impossibility, it should be revised to make compliance possible. It turned out that to accomplish this result every provision in the code had to be substantially rewritten. The final result was, however, a triumph of draftsmanship. It was clear, consistent with itself, and demanded nothing of the subject that did not lie easily within his powers. It was printed and distributed free of charge on every street corner.

However, before the effective date for the new code had arrived, it was discovered that so much time had been spent in successive revisions of Rex's original draft, that the substance of the code had been seriously overtaken by events. Ever since Rex assumed the throne there had been a suspension of ordinary legal processes and this had brought about important economic and institutional changes within the country. Accommodation to these altered conditions required many changes of substance in the law. Accordingly as soon as the new code became legally effective, it was subjected to a daily stream of amendments. Again popular discontent mounted; an anonymous pamphlet appeared on the streets carrying scurrilous cartoons of the king and a leading article with the title: "A law that changes every day is worse than no law at all."

Within a short time this source of discontent began to cure itself as the pace of amendment gradually slackened. Before this had occurred to any noticeable degree, however, Rex announced an important decision. Reflecting on the misadventures of his reign, he concluded that much of the trouble lay in bad advice he had received from experts. He accordingly declared he was reassuming the judicial power in his own person. In this way he could directly control the application of the new code and insure his country against another crisis. He began to spend practically all of his time hearing and deciding cases arising under the new code.

As the king proceeded with this task, it seemed to bring to a belated blossoming his long dormant powers of generalization. His opinions began, indeed, to reveal a confident and almost exuberant virtuosity as he deftly distinguished his own previous decisions, exposed the principles on which he acted, and laid down guidelines for the disposition of future controversies. For Rex's subjects a new day seemed about to dawn when they could finally conform their conduct to a coherent body of rules.

This hope was, however, soon shattered. As the bound volumes of Rex's judgments became available and were subjected to closer study, his subjects were appalled to discover that there existed no discernible relation between those judgments and the code they purported to apply. Insofar as it found expression in the actual disposition of controversies, the new code might just as well not have existed at all. Yet in virtually every one of his decisions Rex declared and redeclared the code to be the basic law of his kingdom.

Leading citizens began to hold private meetings to discuss what measures, short of open revolt, could be taken to get the king away from the bench and back on the throne. While these discussions were going on Rex suddenly died, old before his time and deeply disillusioned with his subjects.

The first act of his successor, Rex II, was to announce that he was taking the powers of government away from the lawyers and placing them in the hands of psychiatrists and experts in public relations. This way, he explained, people could be made happy without rules.

THE CONSEQUENCES OF FAILURE

Rex's bungling career as legislator and judge illustrates that the attempt to create and maintain a system of legal rules may miscarry in at least eight ways; there are in this enterprise, if you will, eight distinct routes to disaster. The first and most obvious lies in a failure to achieve rules at all, so that every issue must be decided on an ad hoc basis. The other routes are: (2) a failure to publicize, or at least to make available to the affected party, the rules he is expected to observe; (3) the abuse of retroactive legislation, which not only cannot itself guide action, but undercuts the integrity of rules prospective in effect, since it puts them under the threat of retrospective change; (4) a failure to make rules understandable; (5) the enactment of contradictory rules or (6) rules that require conduct beyond the powers of the affected party; (7) introducing such frequent changes in the rules that the subject cannot

orient his action by them; and, finally, (8) a failure of congruence between the rules as announced and their actual administration.

A total failure in any one of these eight directions does not simply result in a bad system of law; it results in something that is not properly called a legal system at all, except perhaps in the Pickwickian sense in which a void contract can still be said to be one kind of contract. Certainly there can be no rational ground for asserting that a man can have a moral obligation to obey a legal rule that does not exist, or is kept secret from him, or that came into existence only after he had acted, or was unintelligible, or was contradicted by another rule of the same system, or commanded the impossible, or changed every minute. It may not be impossible for a man to obey a rule that is disregarded by those charged with its administration, but at some point obedience becomes futile—as futile, in fact, as casting a vote that will never be counted. As the sociologist Simmel has observed, there is a kind of reciprocity between government and the citizen with respect to the observance of rules.[1] Government says to the citizen in effect, "These are the rules we expect you to follow. If you follow them, you have our assurance that they are the rules that will be applied to your conduct." When this bond of reciprocity is finally and completely ruptured by government, nothing is left on which to ground the citizen's duty to observe the rules.

The citizen's predicament becomes more difficult when, though there is no total failure in any direction, there is a general and drastic deterioration in legality, such as occurred in Germany under Hitler.[2] A situation begins to develop, for example, in which though some laws are published, others, including the most important, are not. Though most laws are prospective in effect, so free a use is made of retrospective legislation that no law is immune to change ex post facto if it suits the convenience of those in power. For the trial of criminal cases concerned with loyalty to the regime, special military tribunals are established and these tribunals disregard, whenever it suits their convenience, the rules that are supposed to control their decisions. Increasingly the principal object of government seems to be not that of giving the citizen rules by which to shape his conduct, but to frighten him into impotence. As such a situation develops, the problem faced by the citizen is not so simple as that of a voter who knows with certainty that his ballot will not be counted. It is more like that of the voter who knows that the odds are against his ballot being counted at all, and that if it is counted, there is a good chance that it will be counted for the side against which he actually voted. A citizen in this predicament has to decide for himself whether to stay with the system and cast his ballot as a kind of symbolic act expressing the hope of a better day. So it was with the German citizen under Hitler faced with deciding whether he had an obligation to obey such portions of the laws as the Nazi terror had left intact.

In situations like these there can be no simple principle by which to test the citizen's obligation of fidelity to law, any more than there can be such a principle for testing his right to engage in a general revolution. One thing is, however, clear. A mere respect for constituted authority must not be confused with fidelity to law. Rex's subjects, for example, remained faithful to him as king throughout his long and inept reign. They were not faithful to his law, for he never made any. . . .

. . . Do the principles expounded in [the last section] represent some variety of natural law [which holds that law and morality are necessarily linked—EDS.]? The answer is an emphatic, though qualified, yes.

What I have tried to do is to discern and articulate the natural laws of a particular kind of human undertaking, which I have described as "the enterprise of subjecting human conduct to the governance of rules." These natural laws have nothing to do with any "brooding omnipresence in the skies." Nor have they the slightest affinity with any such proposition as that the practice of contraception is a violation of God's law. They remain entirely terrestrial in origin and application. They are not "higher" laws; if any metaphor of elevation is appropriate they should be called "lower" laws. They are like the natural laws of carpentry, or at least those laws respected by a carpenter who wants the house he builds to remain standing and serve the purpose of those who live in it.

> . . . Now clearly the ability to prophesy presupposes order of some sort. . . . If we are to predict intelligently what the courts will do in fact, we must ask what they are trying to do. We must indeed go further and participate vicariously in the whole purposive effort that goes into creating and maintaining a system for directing human conduct by rules. If we are to understand that effort, we must understand that many of its characteristic problems are moral in nature. Thus, we need to put ourselves in the place of the judge faced with a statute extremely vague in its operative terms yet disclosing clearly enough in its preamble an objective the judge considers plainly unwise. We need to share the anguish of the weary legislative draftsman who at 2:00 A.M. says to himself, "I know this has got to be right and if it isn't people may be hauled into court for things we don't mean to cover at all. But how long must I go on rewriting it?"
>
> A concentration on the order imposed by law in abstraction from the purposive effort that goes into creating it is by no means [rare]. Professor Friedmann, for example, in an attempt to offer a neutral concept of law that will not import into the notion of law itself any particular ideal of substantive justice, proposes the following definition:
>
> > the rule of law simply means the "existence of public order." It means organized government, operating through the various instruments and channels of legal command. In this sense, all modern societies live under the rule of law, fascist as well as socialist and liberal states.[3]
>
> Now it is plain that a semblance of "public order" can be created by lawless terror, which may serve to keep people off the streets and in their homes. Obviously, Friedmann does not have this sort of order in mind, for he speaks of the "organized government, operating through the various instruments and channels of legal command." But beyond this vague intimation of the kind of order he has in mind he says nothing. He plainly indicates, however, a conviction that, considered just "as law," the law of Nazi Germany was as much law as that of any other nation. This proposition, I need not say, is completely at odds with the analysis presented here.

NOTES

1. *The Sociology of Georg Simmel* (1950), trans. Wolff, § 4. "Interaction in the Idea of 'Law,' " pp. 186–89; see also Chapter 4, "Subordination under a Principle," pp. 250–67. Simmel's discussion is worthy of study by those concerned with defining the conditions under which the ideal of "the rule of law" can be realized.
2. I have discussed some of the features of this deterioration in my article, "Positivism and Fidelity to Law," 71 *Harvard Law Review* 630, 648–57 (1958). This article makes no attempt at a comprehensive survey of all the postwar judicial decisions in Germany concerned with events occurring during the Hitler regime. Some of the later decisions rested the nullity of judgments rendered by the courts under Hitler not on the ground that the statutes applied were void, but on the ground that the Nazi judges misinterpreted the statutes of their own government. See Pappe, "On the Validity of Judicial Decisions in the Nazi Era," 23 *Modern Law Review* 260–74 (1960). Dr. Pappe makes more of this distinction than seems to me appropriate. After all, the meaning of a statute depends in part on accepted modes of interpretation. Can it be said that the postwar German courts gave full effect to Nazi laws when they interpreted them by their own standards instead of the quite different standards current during the Nazi regime? Moreover, with statutes of the kind involved, filled as they were with vague phrases and unrestricted delegations of power, it seems a little out of place to strain over questions of their proper interpretation.
3. *Law and Social Change* (1951), p. 281.

REVIEW AND DISCUSSION QUESTIONS

1. List the ways Rex failed to make law.
2. In what sense has Fuller shown there is an inherent "morality" of law?
3. How might Fuller be said to provide an account of the ideal of the rule of law?
4. How does law differ from mere maintenance of order, according to Fuller?

The Rule of Law and Its Virtues

Joseph Raz

In this essay, Joseph Raz begins with a discussion of the nature of the ideal of the rule of law itself: What, precisely, does that ideal mean? His answer leads to a discussion of various principles that can be derived from the basic idea of the rule of law. He then describes the values of the rule of law and concludes with a brief critique of the attempt by Lon Fuller to establish a link between law and morality. Joseph Raz is professor of philosophy at Oxford University.

. . . The rule of law is a political ideal which a legal system may lack or may possess to a greater or lesser degree. That much is common ground. It is also to be insisted that the rule of law is just one of the virtues which a legal system may possess and by which it is to be judged. It is not to be confused with democracy, justice, equality (before the law or otherwise), human rights of any kind or respect for persons or for the dignity of man. A nondemocratic legal system, based on the denial of human rights, on extensive poverty, on racial segregation, sexual inequalities, and religious persecution may, in principle, conform to the requirements of the rule of law better than any of the legal systems of the more enlightened Western democracies. This does not mean that it will be better than those Western democracies. It will be an immeasurably worse legal system, but it will excel in one respect: in its conformity to the rule of law.

Given the promiscuous use made in recent years of the expression "the rule of law" it is hardly surprising that my claim will alarm many. We have reached the stage in which no purist can claim that truth is on his side and blame the others for distorting the notion of the rule of law. . . .

1. THE BASIC IDEA

"The rule of law" means literally what it says: the rule of the law. Taken in its broadest sense this means that people should obey the law and be ruled by it. But in political and legal theory it has come to be read in a narrower sense, that the government shall be ruled by the law and subject to it. The ideal of the rule of law in this sense is often expressed by the phrase "government by law and not by men." No sooner does one use these formulas than their obscurity becomes evident. Surely government must be both by law and by men. It is said that the rule of law means that all government action must have foundation in law, must be authorized by law. But is not that a tautology? Actions not authorized by law cannot be the actions of the government as a government. They would be without legal effect and often unlawful. . . .

. . . If government is, by definition, government authorized by law the rule of law seems to amount to an empty tautology, not a political ideal.

The solution to this riddle is in the difference between the professional and the lay sense of "law." For the lawyer anything is the law if it meets the conditions of validity laid down in the system's rules of recognition [i.e., its accepted test of legal validity—EDS.] or in other rules of the system. This includes the constitution, parliamentary legislation, ministerial regulations, policemen's orders, the regulations of limited companies, conditions imposed in trading licenses, etc. To the layman the law consists only of a subclass of these. To him the law is essentially a set of open, general, and relatively stable laws. Government by law and not by men is not a tautology if "law" means general, open, and relatively stable law. In fact, the danger of this interpretation is that the rule of law might set too strict a requirement, one which no legal system can meet and which embodies very little virtue. It is humanly inconceivable that law can consist only of general rules and it is very undesirable that it should. Just as we need government both by laws and by men, so we need both general and particular laws to carry out the jobs for which we need the law.

The doctrine of the rule of law does not deny that every legal system should consist of both

general, open, and stable rules (the popular conception of law) and particular laws (legal orders), an essential tool in the hands of the executive and the judiciary alike. As we shall see, what the doctrine requires is the subjection of particular laws to general, open, and stable ones. It is one of the important principles of the doctrine that *the making of particular laws should be guided by open and relatively stable general rules.*

This principle shows how the slogan of the rule of law and not of men can be read as a meaningful political ideal. The principle does not, however, exhaust the meaning of "the rule of law" and does not by itself illuminate the reasons for its alleged importance. Let us, therefore, return to the literal sense of "the rule of law." It has two aspects: (1) that people should be ruled by the law and obey it, and (2) that the law should be such that people will be able to be guided by it. As was noted above, it is with the second aspect that we are concerned: the law must be capable of being obeyed. A person conforms with the law to the extent that he does not break the law. But he obeys the law only if part of his reason for conforming is his knowledge of the law. Therefore, if the law is to be obeyed *it must be capable of guiding the behaviour of its subjects.* It must be such that they can find out what it is and act on it.

This is the basic intuition from which the doctrine of the rule of law derives: the law must be capable of guiding the behaviour of its subjects. It is evident that this conception of the rule of law is a formal one. It says nothing about how the law is to be made: by tyrants, democratic majorities, or any other way. It says nothing about fundamental rights, about equality, or justice. It may even be thought that this version of the doctrine is formal to the extent that it is almost devoid of content. This is far from the truth. Most of the requirements which were associated with the rule of law before it came to signify all the virtues of the state can be derived from this one basic idea.

2. SOME PRINCIPLES

Many of the principles which can be derived from the basic idea of the rule of law depend for their validity or importance on the particular circumstances of different societies. There is little point in trying to enumerate them all, but some of the more important ones might be mentioned:

(1) All Laws Should Be Prospective, Open, and Clear. One cannot be guided by a retroactive law. It does not exist at the time of action. Sometimes it is then known for certain that a retroactive law will be enacted. When this happens retroactivity does not conflict with the rule of law (though it may be objected to on other grounds). The law must be open and adequately publicized. If it is to guide people they must be able to find out what it is. For the same reason its meaning must be clear. An ambiguous, vague, obscure, or imprecise law is likely to mislead or confuse at least some of those who desire to be guided by it.

(2) Laws Should Be Relatively Stable. They should not be changed too often. If they are frequently changed people will find it difficult to find out what the law is at any given moment and will be constantly in fear that the law has been changed since they last learnt what it was. But more important still is the fact that people need to know the law not only for short-term decisions (where to park one's car, how much alcohol is allowed duty free, etc.) but also for long-term planning. Knowledge of at least the general outlines and sometimes even of details of tax law and company law are often important for business plans which will bear fruit only years later. Stability is essential if people are to be guided by law in their long-term decisions.[1]

. . . [C]onformity to the rule of law is often a matter of degree, not only when the conformity of the legal system as a whole is at stake, but also with respect to single laws. A law is either retroactive or not, but it can be more or less clear, more or less stable, etc. It should be remembered, however, that by asserting that conformity to the principles is a matter of degree, it is not meant that the degree of conformity can be quantitatively measured by counting the number of infringements, or some such method. Some infringements are worse than others. . . .

(3) The Making of Particular Laws (Particular Legal Orders) Should Be Guided by Open, Stable, Clear, and General Rules. . . . Racial, religious,

and all manner of discrimination are not only compatible but often institutionalized by general rules.

The formal conception of the rule of law which I am defending does not object to particular legal orders as long as they are stable, clear, etc. But of course particular legal orders are mostly used by government agencies to introduce flexibility into the law. A police constable regulating traffic, a licensing authority granting a licence under certain conditions, all these and their like are among the more ephemeral parts of the law. As such they run counter to the basic idea of the rule of law. They make it difficult for people to plan ahead on the basis of their knowledge of the law. This difficulty is overcome to a large extent if particular laws of an ephemeral status are enacted only within a framework set by general laws which are more durable and which impose limits on the unpredictability introduced by the particular orders. . . .

(4) The Independence of the Judiciary Must Be Guaranteed. It is of the essence of municipal legal systems that they institute judicial bodies charged, among other things, with the duty of applying the law to cases brought before them and whose judgments and conclusions as to the legal merits of those cases are final. Since just about any matter arising under any law can be subject to a conclusive court judgment, it is obvious that it is futile to guide one's action on the basis of the law if when the matter comes to adjudication the courts will not apply the law and will act for some other reasons. The point can be put even more strongly. Since the court's judgment establishes conclusively what is the law in the case before it, the litigants can be guided by law only if the judges apply the law correctly. Otherwise people will only be able to be guided by their guesses as to what the courts are likely to do—but these guesses will not be based on the law but on other considerations.

The rules concerning the independence of the judiciary—the method of appointing judges, their security of tenure, the way of fixing their salaries, and other conditions of service—are designed to guarantee that they will be free from extraneous pressures and independent of all authority save that of the law. They are, therefore, essential for the preservation of the rule of law.

(5) The Principles of Natural Justice Must Be Observed. Open and fair hearing, absence of bias, and the like are obviously essential for the correct application of the law and thus, through the very same considerations mentioned above, to its ability to guide action.

(6) The Courts Should Have Review Powers over the Implementation of the Other Principles. This includes review of both subordinate and parliamentary legislation and of administrative action, but in itself it is a very limited review—merely to ensure conformity to the rule of law.

(7) The Courts Should Be Easily Accessible. Given the central position of the courts in ensuring the rule of law (see principles 4 and 6) it is obvious that their accessibility is of paramount importance. Long delays, excessive costs, etc., may effectively turn the most enlightened law into a dead letter and frustrate one's ability effectively to guide oneself by the law.

(8) The Discretion of the Crime-Preventing Agencies Should Not Be Allowed to Pervert the Law. Not only the courts but also the actions of the police and the prosecuting authorities can subvert the law. The prosecution should not be allowed, for example, to decide not to prosecute for commission of certain crimes, or for crimes committed by certain classes of offenders. The police should not be allowed to allocate its resources so as to avoid all effort to prevent and detect certain crimes or prosecute certain classes of criminals. . . .

This list is very incomplete.[2] My purpose in listing them was merely to illustrate the power and fruitfulness of the formal conception of the rule of law. It should, however, be remembered that in the final analysis the doctrine rests on its basic idea that the law should be capable of providing effective guidance. The principles do not stand on their own. They must be constantly interpreted in the light of the basic idea.

The eight principles listed fall into two groups. Principles 1 to 3 require that the law should conform to standards designed to enable it effectively to guide action. Principles 4 to 8 are designed to ensure that the legal machinery of enforcing the law should not deprive it of its ability to guide through distorted enforcement and that it shall be

capable of supervising conformity to the rule of law and provide effective remedies in cases of deviation from it. All the principles directly concern the system and method of government in matters directly relevant to the rule of law. Needless to say, many other aspects in the life of a community may, in more indirect ways, either strengthen or weaken the rule of law. A free press run by people anxious to defend the rule of law is of great assistance in preserving it, just as a gagged press or one run by people wishing to undermine the rule of law is a threat to it. But we need not be concerned here with these more indirect influences.

3. THE VALUE OF THE RULE OF LAW

One of the merits of the doctrine of the rule of law I am defending is that there are so many values it does not serve. Conformity to the rule of law is a virtue, but only one of the many virtues a legal system should possess. This makes it all the more important to be clear on the values which the rule of law does serve.

The rule of law is often rightly contrasted with arbitrary power. Arbitrary power is broader than the rule of law. Many forms of arbitrary rule are compatible with the rule of law. A ruler can promote general rules based on whim or self-interest, etc., without offending against the rule of law. But certainly many of the more common manifestations of arbitrary power run foul of the rule of law. A government subjected to the rule of law is prevented from changing the law retroactively or abruptly or secretly whenever this suits its purposes. The one area where the rule of law excludes all forms of arbitrary power is in the law-applying function of the judiciary where the courts are required to be subject only to the law and to conform to fairly strict procedures. No less important is the restraint imposed by the rule of law on the making of particular laws and thus on the powers of the executive. The arbitrary use of power for personal gain, out of vengeance or favouritism, is most commonly manifested in the making of particular legal orders. These possibilities are drastically restricted by close adherence to the rule of law. . . .

. . . Since it is universally believed that it is wrong to use public powers for private ends any such use is in itself an instance of arbitrary use of power. As we have seen the rule of law does help to curb such forms of arbitrary power.

But there are more reasons for valuing the rule of law. We value the ability to choose styles and forms of life, to fix long-term goals and effectively direct one's life towards them. One's ability to do so depends on the existence of stable, secure frameworks for one's life and actions. The law can help to secure such fixed points of reference in two ways: (1) by stabilizing social relationships which but for the law may disintegrate or develop in erratic and unpredictable ways; (2) by a policy of self-restraint designed to make the law itself a stable and safe basis for individual planning. This last aspect is the concern of the rule of law.

This second virtue of the rule of law is often . . . identified as the protection of individual freedom. This is right in the sense of freedom in which it is identified with an effective ability to choose between as many options as possible. Predictability in one's environment does increase one's power of action.[3] If this is freedom, well and good. . . . But it has no bearing on the existence of spheres of activity free from governmental interference and is compatible with gross violations of human rights.

More important than both these considerations is the fact that observance of the rule of law is necessary if the law is to respect human dignity. Respecting human dignity entails treating humans as persons capable of planning and plotting their future. Thus, respecting people's dignity includes respecting their autonomy, their right to control their future. A person's control over his life is never complete. It can be incomplete in any one of several respects. The person may be ignorant of his options, unable to decide what to do, incapable of realizing his choices or frustrated in his attempts to do so, or he may have no choice at all (or at least none which is worth having). All these failures can occur through natural causes or through the limitations of the person's own character and abilities. . . .

The law can violate people's dignity in many ways. Observing the rule of law by no means guarantees that such violations do not occur. But it is clear that deliberate disregard for the rule of law violates human dignity. It is the business of law to guide human action by affecting people's options.

The law may, for example, institute slavery without violating the rule of law. But deliberate violation of the rule of law violates human dignity. The violation of the rule of law can take two forms. It may lead to uncertainty or it may lead to frustrated and disappointed expectations. It leads to the first when the law does not enable people to foresee future developments or to form definite expectations (as in cases of vagueness and most cases of wide discretion). It leads to frustrated expectations when the appearance of stability and certainty which encourages people to rely and plan on the basis of the existing law is shattered by retroactive law-making or by preventing proper law-enforcement, etc. The evils of uncertainty are in providing opportunities for arbitrary power and restricting people's ability to plan for their future. The evils of frustrated expectations are greater. Quite apart from the concrete harm they cause they also offend dignity in expressing disrespect for people's autonomy. The law in such cases encourages autonomous action only in order to frustrate its purpose. When such frustration is the result of human action or the result of the activities of social institutions then it expresses disrespect. Often it is analogous to entrapment: one is encouraged innocently to rely on the law and then that assurance is withdrawn and one's very reliance is turned into a cause of harm to one. A legal system which does in general observe the rule of law treats people as persons at least in the sense that it attempts to guide their behaviour through affecting the circumstances of their action. It thus presupposes that they are rational autonomous creatures and attempts to affect their actions and habits by affecting their deliberations.

Conformity to the rule of law is a matter of degree. Complete conformity is impossible (some vagueness is inescapable) and maximal possible conformity is on the whole undesirable (some controlled administrative discretion is better than none). It is generally agreed that general conformity to the rule of law is to be highly cherished. But one should not take the value of the rule of law on trust nor assert it blindly. Disentangling the various values served by the rule of law helps to assess intelligently what is at stake in various possible or actual violations. Some cases insult human dignity, give free rein to arbitrary power, frustrate one's expectations, and undermine one's ability to plan. Others involve only some of these evils. The evil of different violations of the rule of law is not always the same despite the fact that the doctrine rests on the solid core of its basic idea.

4. THE RULE OF LAW AND ITS ESSENCE

Lon Fuller has claimed that the principles of the rule of law which he enumerated are essential for the existence of law. This claim if true is crucial to our understanding not only of the rule of law but also of the relation of law and morality. I have been treating the rule of law as an ideal, as a standard to which the law ought to conform but which it can and sometimes does violate most radically and systematically. Fuller, while allowing that deviations from the ideal of the rule of law can occur, denies that they can be radical or total. A legal system must of necessity conform to the rule of law to a certain degree, he claims. From this claim he concludes that there is an essential link between law and morality. Law is necessarily moral, at least in some respects. . . .

Clearly, the extent to which generality, clarity, prospectivity, etc., are essential to the law is minimal and is consistent with gross violations of the rule of law. But are not considerations of the kind mentioned sufficient to establish that there is necessarily at least some moral value in every legal system? I think not. The rule of law is essentially a negative value. The law inevitably creates a great danger of arbitrary power—the rule of law is designed to minimize the danger created by the law itself. Similarly, the law may be unstable, obscure, retrospective, etc., and thus infringe people's freedom and dignity. The rule of law is designed to prevent this danger as well. . . . A person who cannot communicate cannot claim any moral merit for being honest. A person who through ignorance or inability cannot kill another by poison deserves no credit for it. Similarly, that the law cannot sanction arbitrary force or violations of freedom and dignity through total absence of generality, prospectivity, or clarity is no moral credit to the law. It only means that there are some kinds of evil which cannot be brought about by the law. But this is no virtue in the law just as it is no virtue in the law that

it cannot rape or murder (all it can do is sanction such actions).

Fuller's attempt to establish a necessary connection between law and morality fails. In so far as conformity to the rule of law is a moral virtue it is an ideal which should but may fail to become a reality. There is another argument, however, which establishes an essential connection between the law and the rule of law, though it does not guarantee any virtue to the law. Conformity to the rule of law is essential for securing whatever purposes the law is designed to achieve. This statement should be qualified. We could divide the purposes a law is intended to serve into two kinds: those which are secured by conformity with the law in itself and those further consequences of conformity with the law or of knowledge of its existence which the law is intended to secure. Thus a law prohibiting racial discrimination in government employment has as its direct purpose the establishment of racial equality in the hiring, promotion, and conditions of service of government employees (since discriminatory action is a breach of law). Its indirect purposes may well be to improve race relations in the country in general, prevent a threat of a strike by some trade unions, or halt the decline in popularity of the government.

Conformity to the rule of law does not always facilitate realization of the indirect purposes of the law, but it is essential to the realization of its direct purposes. These are achieved by conformity with the law which is secured (unless accidentally) by people taking note of the law and guiding themselves accordingly. Therefore, if the direct purposes of the law are not to be frustrated it must be capable of guiding human behaviour, and the more it conforms to the principles of the rule of law the better it can do so.

In section 3 we saw that conformity to the rule of law is one among many moral virtues which the law should possess. The present consideration shows that the rule of law is not merely a moral virtue—it is a necessary condition for the law to be serving directly any good purpose at all. Of course, conformity to the rule of law also enables the law to serve bad purposes. That does not show that it is not a virtue, just as the fact that a sharp knife can be used to harm does not show that being sharp is not a good-making characteristic for knives. At most it shows that from the point of view of the present consideration it is not a moral good. Being sharp is an inherent good-making characteristic of knives. A good knife is, among other things, a sharp knife. Similarly, conformity to the rule of law is an inherent value of laws, indeed it is their most important inherent value. It is of the essence of law to guide behaviour through rules and courts in charge of their application. Therefore, the rule of law is the specific excellence of the law. Since conformity to the rule of law is the virtue of law in itself, law as law regardless of the purposes it serves, it is understandable and right that the rule of law is thought of as among the few virtues of law which are the special responsibility of the courts and the legal profession.

Regarding the rule of law as the inherent or specific virtue of law is a result of an instrumental conception of law. The law is not just a fact of life. It is a form of social organization which should be used properly and for the proper ends. It is a tool in the hands of men differing from many others in being versatile and capable of being used for a large variety of proper purposes. As with some other tools, machines, and instruments a thing is not of the kind unless it has at least some ability to perform its function. A knife is not a knife unless it has some ability to cut. The law to be law must be capable of guiding behaviour, however inefficiently. Like other instruments, the law has a specific virtue which is morally neutral in being neutral as to the end to which the instrument is put. It is the virtue of efficiency; the virtue of the instrument as an instrument. For the law this virtue is the rule of law. Thus the rule of law is an inherent virtue of the law, but not a moral virtue as such.

The special status of the rule of law does not mean that conformity with it is of no moral importance. Quite apart from the fact that conformity to the rule of law is also a moral virtue, it is a moral requirement when necessary to enable the law to perform useful social functions; just as it may be of moral importance to produce a sharp knife when it is required for a moral purpose. In the case of the rule of law this means that it is virtually always of great moral value. . . .

NOTES

1. Of course, uncertainty generated by instability of law also affects people's planning of action. If it did not, stability would not have any impact either. The point is that only if the law is stable are people guided by *their knowledge of the content of the law.*
2. For a list similar to mine see Lon Fuller's *The Morality of Law*, 2nd ed., ch. 2. His discussion of many of the principles is full of good sense. My main reason for abandoning some of his principles is a difference of views on conflicts between the laws of one system.
3. But then welfare law and governmental manipulation of the economy also increase freedom by increasing—if successful—people's welfare. If the rule of law is defended as the bulwark of freedom in this sense, it can hardly be used to oppose in principle governmental management of the economy.

REVIEW AND DISCUSSION QUESTIONS

1. What is the basic idea of the rule of law, according to Raz?
2. Describe the different principles that Raz thinks are implicit in the idea of the rule of law.
3. What are the values that Raz argues the rule of law serves? What values does it not guarantee?
4. Describe why Raz rejects Fuller's claim that there is a necessary connection between law and morality.
5. Using Raz's metaphor that the law should be a sharp knife, discuss how he understands the importance of conformity to the rule of law and the importance of obeying law.

Grudge Informers and the Rule of Law

H. L. A. Hart

Using the example of people who used Nazi law and power to work out grudges against other citizens, H. L. A. Hart continues the consideration of a range of issues that will occupy us in later sections. These issues include the duty to obey law, the relationship between law and morality, and the ideal of the rule of law. H. L. A. Hart was professor of jurisprudence at Oxford University.

Gustav Radbruch was a German legal philosopher who had shared the "positivist" doctrine [that law and morality are in principle distinct] until the Nazi tyranny. Prior to his recantation of positivism, he held that resistance to law was a matter for the personal conscience, to be thought out by the individual as a moral problem, and the validity of a law could not be disproved by showing that its requirements were morally evil or even by showing that the effect of compliance with the law would be more evil than the effect of disobedience. . . .

Radbruch, however, had concluded from the ease with which the Nazi regime had exploited subservience to mere law—as expressed, as he thought, in the "positivist" slogan "law is law"—and from the failure of the German legal profession to protest against the enormities which they were required to perpetrate in the name of law, that "positivism" (meaning here the insistence on the separation of law as it is from law as it ought to be) had powerfully contributed to the horrors.

After the war Radbruch's conception of law as containing in itself the essential moral principle of humanitarianism was applied in practice by German courts in certain cases in which local war criminals, spies, and informers under the Nazi regime were punished. The special importance of these cases is that the persons accused of these crimes claimed that what they had done was not illegal under the laws of the regime in force at the

time these actions were performed. This plea was met with the reply that the laws upon which they relied were invalid as contravening the fundamental principles of morality. Let me cite briefly one of these cases.

In 1944 a woman, wishing to be rid of her husband, denounced him to the authorities for insulting remarks he had made about Hitler while home on leave from the German army. The wife was under no legal duty to report his acts, though what he had said was apparently in violation of statutes making it illegal to make statements detrimental to the government of the Third Reich or to impair by any means the military defense of the German people. The husband was arrested and sentenced to death, apparently pursuant to these statutes, though he was not executed but was sent to the front. In 1949 the wife was prosecuted in a West German court for an offense which we would describe as illegally depriving a person of his freedom. This was punishable as a crime under the German Criminal Code of 1871, which had remained in force continuously since its enactment. The wife pleaded that her husband's imprisonment was pursuant to the Nazi statutes and hence that she had committed no crime. The court of appeal to which the case ultimately came held that the wife was guilty of procuring the deprivation of her husband's liberty by denouncing him to the German courts, even though he had been sentenced by a court for having violated a statute, since, to quote the words of the court, the statute "was contrary to the sound conscience and sense of justice of all decent human beings." This reasoning was followed in many cases which have been hailed as a triumph of the doctrines of natural law and as signaling the overthrow of positivism. The unqualified satisfaction with this result seems to me to be hysteria. Many of us might applaud the objective—that of punishing a woman for an outrageously immoral act—but this was secured only by declaring a statute established since 1934 not to have the force of law, and at least the wisdom of this course must be doubted. There were, of course, two other choices. One was to let the woman go unpunished; one can sympathize with and endorse the view that this might have been a bad thing to do. The other was to face the fact that if the woman were to be punished it must be pursuant to the introduction of a frankly retrospective law and with a full consciousness of what was sacrificed in securing her punishment in this way. Odious as retrospective criminal legislation and punishment may be, to have pursued it openly in this case would at least have had the merits of candour. It would have made plain that in punishing the woman a choice had to be made between two evils, that of leaving her unpunished and that of sacrificing a very precious principle of morality endorsed by most legal systems. Surely if we have learned anything from the history of morals it is that the thing to do with a moral quandary is not to hide it. Like nettles, the occasions when life forces us to choose between the lesser of two evils must be grasped with the consciousness that they are what they are.

REVIEW AND DISCUSSION QUESTIONS

1. Describe the facts relating to the woman accused of being a grudge informer.
2. What solution to the problem does Hart recommend? Why does he recommend it?
3. Explain why Hart thinks that retrospective legislation, whatever its disadvantages, is better than the alternative(s).
4. Do you think that "positivism" was a relevant factor in the German experience, as Radbruch did? Explain.

The Problem of the Grudge Informer

Lon Fuller

In the following essay, Lon Fuller discusses an issue that is often confronted as societies undergo rapid political change. The specific question he poses, whether to prosecute grudge informers from a prior regime, raises complex questions about the meaning and importance of the rule of law and about the nature and purpose of law itself. Lon Fuller was professor of law at Harvard Law School.

By a narrow margin you have been elected Minister of Justice of your country, a nation of some twenty million inhabitants. At the outset of your term of office you are confronted by a serious problem that will be described below. But first the background of this problem must be presented.

For many decades your country enjoyed a peaceful, constitutional and democratic government. However, some time ago it came upon bad times. Normal relations were disrupted by a deepening economic depression and by an increasing antagonism among various factional groups, formed along economic, political, and religious lines. The proverbial man on horseback appeared in the form of the Headman of a political party or society that called itself the Purple Shirts.

In a national election attended by much disorder the Headman was elected President of the Republic and his party obtained a majority of the seats in the General Assembly. The success of the party at the polls was partly brought about by a campaign of reckless promises and ingenious falsifications, and partly by the physical intimidation of nightriding Purple Shirts who frightened many people away from the polls who would have voted against the party.

When the Purple Shirts arrived in power they took no steps to repeal the ancient Constitution or any of its provisions. They also left intact the Civil and Criminal Codes and the Code of Procedure. No official action was taken to dismiss any government official or to remove any judge from the bench. Elections continued to be held at intervals and ballots were counted with apparent honesty. Nevertheless, the country lived under a reign of terror.

Judges who rendered decisions contrary to the wishes of the party were beaten and murdered. The accepted meaning of the Criminal Code was perverted to place political opponents in jail. Secret statutes were passed, the contents of which were known only to the upper levels of the party hierarchy. Retroactive statutes were enacted which made acts criminal that were legally innocent when committed. No attention was paid by the government to the restraints of the Constitution, of antecedent laws, or even of its own laws. All opposing political parties were disbanded. Thousands of political opponents were put to death, either methodically in prisons or in sporadic night forays of terror. A general amnesty was declared in favor of persons under sentence for acts "committed in defending the fatherland against subversion." Under this amnesty a general liberation of all prisoners who were members of the Purple Shirt party was effected. No one not a member of the party was released under the amnesty.

The Purple Shirts as a matter of deliberate policy preserved an element of flexibility in their operations by acting at times through the party "in the streets," and by acting at other times through the apparatus of the state which they controlled. Choice between the two methods of proceeding was purely a matter of expediency. For example, when the inner circle of the party decided to ruin all the former Socialist-Republicans (whose party put up a last-ditch resistance to the new regime), a dispute arose as to the best way of confiscating their property. One faction, perhaps still influenced by prerevolutionary conceptions, wanted to accomplish this by a statute declaring their goods forfeited for criminal acts. Another wanted to do it by compelling the owners to deed their property over at the point of a bayonet. This group argued against the proposed statute on the ground that it would attract unfavorable comment abroad. The Headman decided in favor of

direct action through the party to be followed by a secret statute ratifying the party's action and confirming the titles obtained by threats of physical violence.

The Purple Shirts have now been overthrown and a democratic and constitutional government restored. Some difficult problems have, however, been left behind by the deposed regime. These you and your associates in the new government must find some way of solving. One of these problems is that of the "grudge informer."

During the Purple Shirt regime a great many people worked off grudges by reporting their enemies to the party or to the government authorities. The activities reported were such things as the private expression of views critical of the government, listening to foreign radio broadcasts, associating with known wreckers and hooligans, hoarding more than the permitted amount of dried eggs, failing to report a loss of identification papers within five days, etc. As things then stood with the administration of justice, any of these acts, if proved, could lead to a sentence of death. In some cases this sentence was authorized by "emergency" statutes; in others it was imposed without statutory warrant, though by judges duly appointed to their offices.

After the overthrow of the Purple Shirts, a strong public demand grew up that these grudge informers be punished. The interim government, which preceded that with which you are associated, temporized on this matter. Meanwhile it has become a burning issue and a decision concerning it can no longer be postponed. Accordingly, your first act as Minister of Justice has been to address yourself to it. You have asked your five Deputies to give thought to the matter and to bring their recommendations to conference. At the conference the five Deputies speak in turn as follows:

First Deputy: "It is perfectly clear to me that we can do nothing about these so-called grudge informers. The acts they reported were unlawful according to the rules of the government then in actual control of the nation's affairs. The sentences imposed on their victims were rendered in accordance with principles of law then obtaining. These principles differed from those familiar to us in ways that we consider detestable. Nevertheless they were then the law of the land. One of the principal differences between that law and our own lies in the much wider discretion it accorded to the judge in criminal matters. This rule and its consequences are as much entitled to respect by us as the reform which the Purple Shirts introduced into the law of wills, whereby only two witnesses were required instead of three. It is immaterial that the rule granting the judge a more or less uncontrolled discretion in criminal cases was never formally enacted but was a matter of tacit acceptance. Exactly the same thing can be said of the opposite rule which we accept that restricts the judge's discretion narrowly. The difference between ourselves and the Purple Shirts is not that theirs was an unlawful government—a contradiction in terms—but lies rather in the field of ideology. No one has a greater abhorrence than I for Purple Shirtism. Yet the fundamental difference between our philosophy and theirs is that we permit and tolerate differences in viewpoint, while they attempted to impose their monolithic code on everyone. Our whole system of government assumes that law is a flexible thing, capable of expressing and effectuating many different aims. The cardinal point of our creed is that when an objective has been duly incorporated into a law or judicial decree it must be provisionally accepted even by those that hate it, who must await their chance at the polls, or in another litigation, to secure a legal recognition for their own aims. The Purple Shirts, on the other hand, simply disregarded laws that incorporated objectives of which they did not approve, not even considering it worth the effort involved to repeal them. If we now seek to unscramble the acts of the Purple Shirt regime, declaring this judgment invalid, that statute void, this sentence excessive, we shall be doing exactly the thing we most condemn in them. I recognize that it will take courage to carry through with the program I recommend and we shall have to resist strong pressures of public opinion. We shall also have to be prepared to prevent the people from taking the law into their own hands. In the long run, however, I believe the course I recommend is the only one that will insure the triumph of the conceptions of law and government in which we believe."

Second Deputy: "Curiously, I arrive at the same conclusion as my colleague, by an exactly opposite route. To me it seems absurd to call the Purple

Shirt regime a lawful government. A legal system does not exist simply because policemen continue to patrol the streets and wear uniforms or because a constitution and code are left on the shelf unrepealed. A legal system presupposes laws that are known, or can be known, by those subject to them. It presupposes some uniformity of action and that like cases will be given like treatment. It presupposes the absence of some lawless power, like the Purple Shirt Party, standing above the government and able at any time to interfere with the administration of justice whenever it does not function according to the whims of that power. All of these presuppositions enter into the very conception of an order of law and have nothing to do with political and economic ideologies. In my opinion law in any ordinary sense of the word ceased to exist when the Purple Shirts came to power. During their regime we had, in effect, an interregnum in the rule of law. Instead of a government of laws we had a war of all against all conducted behind barred doors, in dark alleyways, in palace intrigues, and prisonyard conspiracies. The acts of these so-called grudge informers were just one phase of that war. For us to condemn these acts as criminal would involve as much incongruity as if we were to attempt to apply juristic conceptions to the struggle for existence that goes on in the jungle or beneath the surface of the sea. We must put this whole dark, lawless chapter of our history behind us like a bad dream. If we stir among its hatreds, we shall bring upon ourselves something of its evil spirit and risk infection from its miasmas. I therefore say with my colleague, let bygones be bygones. Let us do nothing about the so-called grudge informers. What they did do was neither lawful nor contrary to law, for they lived, not under a regime of law, but under one of anarchy and terror."

Third Deputy: "I have a profound suspicion of any kind of reasoning that proceeds by an 'either-or' alternative. I do not think we need to assume either, on the one hand, that in some manner the whole of the Purple Shirt regime was outside the realm of law, or, on the other, that all of its doings are entitled to full credence as the act of a lawful government. My two colleagues have unwittingly delivered powerful arguments against these extreme assumptions by demonstrating that both of them lead to the same absurd conclusion, a conclusion that is ethically and politically impossible. If one reflects about the matter without emotion it becomes clear that we did not have during the Purple Shirt regime a 'war of all against all.' Under the surface much of what we call normal human life went on—marriages were contracted, goods were sold, wills were drafted and executed. This life was attended by the usual dislocations—automobile accidents, bankruptcies, unwitnessed wills, defamatory misprints in the newspapers. Much of this normal life and most of these equally normal dislocations of it were unaffected by the Purple Shirt ideology. The legal questions that arose in this area were handled by the courts much as they had been formerly and much as they are being handled today. It would invite an intolerable chaos if we were to declare everything that happened under the Purple Shirts to be without legal basis. On the other hand, we certainly cannot say that the murders committed in the streets by members of the party acting under orders from the Headman were lawful simply because the party had achieved control of the government and its chief had become President of the Republic. If we must condemn the criminal acts of the party and its members, it would seem absurd to uphold every act which happened to be canalized through the apparatus of a government that had become, in effect, the alter ego of the Purple Shirt Party. We must therefore, in this situation, as in most human affairs, discriminate. Where the Purple Shirt philosophy introduced itself and perverted the administration of justice from its normal aims and uses, there we must interfere. Among these perversions of justice I would count, for example, the case of a man who was in love with another man's wife and brought about the death of the husband by informing against him for a wholly trivial offense, that is, for not reporting a loss of his identification papers within five days. This informer was a murderer under the Criminal Code which was in effect at the time of his act and which the Purple Shirts had not repealed. He encompassed the death of one who stood in the way of his illicit passions and utilized the courts for the realization of his murderous intent. He knew that the courts were themselves the pliant instruments of whatever policy the Purple Shirts

might for the moment consider expedient. There are other cases that are equally clear. I admit that there are also some that are less clear. We shall be embarrassed, for example, by the cases of mere busybodies who reported to the authorities everything that looked suspect. Some of these persons acted not from desire to get rid of those they accused, but with a desire to curry favor with the party, to divert suspicions (perhaps ill-founded) raised against themselves, or through sheer officiousness. I don't know how these cases should be handled, and make no recommendation with regard to them. But the fact that these troublesome cases exist should not deter us from acting at once in the cases that are clear, of which there are far too many to permit us to disregard them."

Fourth Deputy: "Like my colleague I too distrust 'either-or' reasoning, but I think we need to reflect more than he has about where we are headed. This proposal to pick and choose among the acts of the deposed regime is thoroughly objectionable. It is, in fact, Purple Shirtism itself, pure and simple. We like this law, so let us enforce it. We like this judgment, let it stand. This law we don't like, therefore it never was a law at all. This governmental act we disapprove, let it be deemed a nullity. If we proceed this way, we take toward the laws and acts of the Purple Shirt government precisely the unprincipled attitude they took toward the laws and acts of the government they supplanted. We shall have chaos, with every judge and every prosecuting attorney a law unto himself. Instead of ending the abuses of the Purple Shirt regime, my colleague's proposal would perpetuate them. There is only one way of dealing with this problem that is compatible with our philosophy of law and government and that is to deal with it by duly enacted law, I mean, by a special statute directed toward it. Let us study this whole problem of the grudge informer, get all the relevant facts, and draft a comprehensive law dealing with it. We shall not then be twisting old laws to purposes for which they were never intended. We shall furthermore provide penalties appropriate to the offense and not treat every informer as a murderer simply because the one he informed against was ultimately executed. I admit that we shall encounter some difficult problems of draftsmanship. Among other things, we shall have to assign a definite legal meaning to 'grudge' and that will not be easy. We should not be deterred by these difficulties, however, from adopting the only course that will lead us out of a condition of lawless, personal rule."

Fifth Deputy: "I find a considerable irony in the last proposal. It speaks of putting a definite end to the abuses of the Purple Shirtism, yet it proposes to do this by resorting to one of the most hated devices of the Purple Shirt regime, the expost facto criminal statute. My colleague dreads the confusion that will result if we attempt without a statute to undo and redress 'wrong' acts of the departed order, while we uphold and enforce its 'right' acts. Yet he seems not to realize that his proposed statute is a wholly specious cure for this uncertainty. It is easy to make a plausible argument for an undrafted statute; we all agree it would be nice to have things down in black and white on paper. But just what would this statute provide? One of my colleagues speaks of someone who had failed for five days to report a loss of his identification papers. My colleague implies that the judicial sentence imposed for that offense, namely death, was so utterly disproportionate as to be clearly wrong. But we must remember that at that time the underground movement against the Purple Shirts was mounting in intensity and that the Purple Shirts were being harassed constantly by people with false identification papers. From their point of view they had a real problem, and the only objection we can make to their solution of it (other than the fact that we didn't want them to solve it) was that they acted with somewhat more rigor than the occasion seemed to demand. How will my colleague deal with this case in his statute, and with all of its cousins and second cousins? Will he deny the existence of any need for law and order under the Purple Shirt regime? I will not go further into the difficulties involved in drafting this proposed statute, since they are evident enough to anyone who reflects. I shall instead turn to my own solution. It has been said on very respectable authority that the main purpose of the criminal law is to give an outlet to the human instinct for revenge. There are times, and I believe this is one of them, when we should allow that instinct to express itself directly without the intervention

of forms of law. This matter of the grudge informers is already in process of straightening itself out. One reads almost every day that a former lackey of the Purple Shirt regime has met his just reward in some unguarded spot. The people are quietly handling this thing in their own way and if we leave them alone, and instruct our public prosecutors to do the same, there will soon be no problem left for us to solve. There will be some disorders, of course, and a few innocent heads will be broken. But our government and our legal system will not be involved in the affair and we shall not find ourselves hopelessly bogged down in an attempt to unscramble all the deeds and misdeeds of the Purple Shirts."

As Minister of Justice which of these recommendations would you adopt?

REVIEW AND DISCUSSION QUESTIONS

1. Describe the five alternatives that the various deputies defend.
2. Fuller's dialogue raises two issues. The first is whether or not a state may legitimately punish someone when what he or she did was wrong, though not illegal. How would you answer that question?
3. The article also raises questions about the validity of laws during the reign of the Purple Shirts. In light of your own understanding of what it is that makes a law valid, and the deputy's arguments, is the law of the Purple Shirts valid? Were laws from the previous constitutional regime still valid? How does that question bear on the choice between the five deputies' positions?
4. How would you answer the question raised at the end of the article? Explain why you reject the advice of each deputy whose position you think is mistaken.
5. Fuller's article shows that this is a dilemma in which each position leads to the next, and the final back to the first. Explain.

Detention of Enemy Combatants

Hamdi v. Rumsfeld

Yaser Hamdi is an American citizen captured by U.S. military forces in Afghanistan. Declaring him an "enemy combatant," the government kept him in custody without access to counsel or the opportunity to challenge his detention. His father filed a writ of *habeas corpus* on his behalf, arguing that Hamdi was being illegally detained and asserting that he was a relief worker who had no military training and merely happened to be in the wrong place at the wrong time.

The case worked its way to a divided Supreme Court. Justice O'Connor wrote the plurality opinion, which holds that Congress authorized the detention of enemy combatants but that due process demands that Hamdi be granted a meaningful opportunity to contest the factual basis for his detention before a neutral decision maker. Concurring in part and dissenting in part, Justice Souter argues that Congress did not authorize the detention of citizens and, thus, the government's holding Hamdi violates the law against imprisoning or detaining citizens except pursuant to an act of Congress. In his dissent, Justice Scalia argues that when our government accuses a citizen of waging war against it, it must prosecute the person—unless Congress has suspended the writ of *habeas corpus.* Because *habeas corpus* has not been suspended, Hamdi must be either charged or released. In a separate dissent, Justice Thomas argues that when it comes to matters of national security, the courts should not second-guess the decisions of the executive branch.

542 U.S. 507 (2004).

JUSTICE O'CONNOR announced the judgment of the Court and delivered an opinion, in which THE CHIEF JUSTICE, JUSTICE KENNEDY, and JUSTICE BREYER join: At this difficult time in our Nation's history, we are called upon to consider the legality of the Government's detention of a United States citizen on United States soil as an "enemy combatant" and to address the process that is constitutionally owed to one who seeks to challenge his classification as such. The United States Court of Appeals for the Fourth Circuit held that petitioner's detention was legally authorized and that he was entitled to no further opportunity to challenge his enemy-combatant label. We now vacate and remand. We hold that although Congress authorized the detention of combatants in the narrow circumstances alleged here, due process demands that a citizen held in the United States as an enemy combatant be given a meaningful opportunity to contest the factual basis for that detention before a neutral decisionmaker.

I

On September 11, 2001, the al Qaeda terrorist network used hijacked commercial airliners to attack prominent targets in the United States. Approximately 3,000 people were killed in those attacks. One week later, in response to those "acts of treacherous violence," Congress passed a resolution authorizing the President to "use all necessary and appropriate force against those nations, organizations, or persons he determines planned, authorized, committed, or aided the terrorist attacks" or "harbored such organizations or persons, in order to prevent any future acts of international terrorism against the United States by such nations, organizations or persons." Authorization for Use of Military Force ("the AUMF"), 115 Stat. 224. Soon thereafter, the President ordered United States Armed Forces to Afghanistan, with a mission to subdue al Qaeda and quell the Taliban regime that was known to support it.

This case arises out of the detention of a man whom the Government alleges took up arms with the Taliban during this conflict. His name is Yaser Esam Hamdi. Born an American citizen in Louisiana in 1980, Hamdi moved with his family to Saudi Arabia as a child. By 2001, the parties agree, he resided in Afghanistan. At some point that year, he was seized by members of the Northern Alliance, a coalition of military groups opposed to the Taliban government, and eventually was turned over to the United States military. . . . The Government contends that Hamdi is an "enemy combatant," and that this status justifies holding him in the United States indefinitely—without formal charges or proceedings—unless and until it makes the determination that access to counsel or further process in warranted.

In June 2002, Hamdi's father, Esam Fouad Hamdi, filed the present petition for a writ of habeas corpus. The elder Hamdi alleges in the petition that he has had no contact with his son since the Government took custody of him in 2001, and that the Government has held his son "without access to legal counsel or notice of any charges pending against him." The petition contends that Hamdi's detention was not legally authorized. It argues that, "[a]s an American citizen, . . . Hamdi enjoys the full protections of the Constitution." . . . Although his habeas petition provides no details with regard to the factual circumstances surrounding his son's capture and detention, Hamdi's father has asserted in documents found elsewhere in the record that this son went to Afghanistan to do "relief work," and that he had been in that country less than two months before September 11, 2001, and could not have received military training. The 20-year-old was travelling on his own for the first time, his father says, and "[b]ecause of his lack of experience, he was trapped in Afghanistan once that military campaign began."

The District Court . . . appointed the federal public defender as counsel for the petitioners, and ordered that counsel be given access to Hamdi. The United States Court of Appeals for the Fourth Circuit reversed that order, holding that the District Court had failed to extend appropriate deference to the Government's security and intelligence interests. . . . It opined that "if Hamdi is indeed an 'enemy combatant' who has captured during hostilities in Afghanistan, the government's present detention of him is a lawful one."

On remand, the Government filed a response and a motion to dismiss the petition. It attached to

its response a declaration from one Michael Mobbs (hereinafter "Mobbs Declaration"), who identified himself as Special Advisor to the Under Secretary of Defense for Policy. . . .

Mobbs then set forth what remains the sole evidentiary support that the Government has provided to the courts for Hamdi's detention. The declaration states that Hamdi "traveled to Afghanistan" in July or August 2001, and that he thereafter "affiliated with a Taliban military unit and received weapons training." It asserts that Hamdi . . . was labeled an enemy combatant "[b]ased upon his interviews and in light of his association with the Taliban." According to the declaration, a series of "U.S. military screening team[s]" determined that Hamdi met "the criteria for enemy combatants," and "a subsequent interview of Hamdi has confirmed that he surrendered and gave his firearm to Northern Alliance forces, which supports his classification as an enemy combatant."

After the Government submitted this declaration, the Fourth Circuit directed the District Court to proceed in accordance with its earlier ruling and, specifically, to "consider the sufficiency of the Mobbs Declaration as an independent matter before proceeding further." The District Court found that the Mobbs Declaration fell "far short" of supporting Hamdi's detention. It criticized the generic and hearsay nature of the affidavit, calling it "little more than the government's 'say-so.' " It ordered the Government to turn over numerous materials for *in camera* review, including copies of all of Hamdi's statements and the notes taken from interviews with him; . . . a list of all interrogators who had questioned Hamdi and their names and addresses; statements by members of the Northern Alliance regarding Hamdi's surrender and capture; a list of the dates and locations of his capture and subsequent detentions; and the names and titles of the United States Government officials who made the determinations that Hamdi was an enemy combatant and that he should be moved to a naval brig. The court indicated that all of these materials were necessary for "meaningful judicial review" of whether Hamdi's detention was legally authorized . . . and whether Hamdi had received sufficient process to satisfy the Due Process Clause of the Constitution and relevant treaties or military regulations.

The Government sought to appeal the production order. . . . The Fourth Circuit reversed. . . . Concluding that the factual averments in the Mobbs Declaration, "if accurate," provided a sufficient basis upon which to conclude that the President had constitutionally detained Hamdi pursuant to the President's war powers, it ordered the habeas petition dismissed. The Fourth Circuit emphasized that the "vital purposes" of the detention of uncharged enemy combatants—preventing those combatants from rejoining the enemy while relieving the military of the burden of litigating the circumstances of wartime captures halfway around the globe—were interests "directly derived from the war powers of Articles I and II." . . .

We now vacate the judgment below and remand.

II

The threshold question before us is whether the Executive has the authority to detain citizens who qualify as "enemy combatants." There is some debate as to the proper scope of this term, and the Government has never provided any court with the full criteria that it uses in classifying individuals as such. It has made clear, however, that, for purposes of this case, the "enemy combatant" that it is seeking to detain is an individual who, it alleges, was "part of or supporting forces hostile to the United States or coalition partners" in Afghanistan and who "engaged in an armed conflict against the United States" there. We therefore answer only the narrow question before us: whether the detention of citizens falling within that definition is authorized. . . .

[W]e agree . . . that Congress has in fact authorized Hamdi's detention, through the AUMF. . . . The AUMF authorizes the President to use "all necessary and appropriate force" against "nations, organizations, or persons" associated with the September 11, 2001, terrorist attacks. There can be no doubt that individuals who fought against the United States in Afghanistan as part of the Taliban, an organization known to have supported the al Qaeda terrorist network responsible for those attacks, are individuals Congress sought to target in passing the AUMF. We conclude that detention of individuals falling into the limited

category we are considering, for the duration of the particular conflict in which they were captured, is so fundamental and accepted an incident to war as to be an exercise of the "necessary and appropriate force" Congress has authorized the President to use. . . .

There is no bar to this Nation's holding one of its own citizens as an enemy combatant. . . .

Hamdi objects, nevertheless, that Congress has not authorized the *indefinite* detention to which he is now subject. . . . The prospect Hamdi raises is not far-fetched. If the Government does not consider this unconventional war won for two generations, and if it maintains during that time that Hamdi might, if released, rejoin forces fighting against the United States, then the position it has taken throughout the litigation of this case suggests that Hamdi's detention could last for the rest of his life.

It is a clearly established principle of the law of war that detention may last no longer than active hostilities. . . . Active combat operations against Taliban fighters apparently are ongoing in Afghanistan. . . . The United States may detain, for the duration of these hostilities, individuals legitimately determined to be Taliban combatants who "engaged in an armed conflict against the United States." If the record establishes that United States troops are still involved in active combat in Afghanistan, those detentions are part of the exercise of "necessary and appropriate force," and therefore are authorized by the AUMF. . . .

III

Even in cases in which the detention of enemy combatants is legally authorized, there remains the question of what process is constitutionally due to a citizen who disputes his enemy-combatant status. . . .

A

Though they reach radically different conclusions on the process that ought to attend the present proceeding, the parties begin on common ground. All agree that, absent suspension, the writ of habeas corpus remains available to every individual detained within the United States. . . . Only in the rarest of circumstances has Congress seen fit to suspend the writ. . . . At all other times, it has remained a critical check on the Executive, ensuring that it does not detain individuals except in accordance with law. . . . All agree suspension of the writ has not occurred here. Thus, it is undisputed that Hamdi was properly before an Article III court to challenge his detention under 28 U.S.C. §2241. . . . Most notably, §2243 provides that "the person detained may, under oath, deny any of the facts set forth in the return or allege any other material facts," and §2246 allows the taking of evidence in habeas proceedings by deposition, affidavit, or interrogatories.

The simple outline of §2241 makes clear both that Congress envisioned that habeas petitioners would have some opportunity to present and rebut facts and that courts in cases like this retain some ability to vary the ways in which they do so as mandated by due process. The Government recognizes the basic procedural protections required by the habeas statute, but asks us to hold that, given both the flexibility of the habeas mechanism and the circumstances presented in this case, the presentation of the Mobbs Declaration to the habeas court completed the required factual development. It suggests two separate reasons for its position that no further process is due.

B

First, the Government urges the adoption of the Fourth Circuit's holding below—that because it is "undisputed" that Hamdi's seizure took place in a combat zone, the habeas determination can be made purely as a matter of law, with no further hearing or factfinding necessary. This argument is easily rejected. . . . [T]he circumstances surrounding Hamdi's seizure cannot in any way be characterized as "undisputed" . . .

C

The Government's second argument requires closer consideration. This is the argument that further factual exploration is unwarranted and inappropriate in light of the extraordinary constitutional interests

at stake. . . . In response, Hamdi emphasizes that this Court consistently has recognized that an individual challenging his detention may not be held at the will of the Executive without recourse to some proceeding before a neutral tribunal to determine whether the Executive's asserted justifications for that detention have basis in fact and warrant in law. . . .

Both of these positions highlight legitimate concerns. And both emphasize the tension that often exists between the autonomy that the Government asserts is necessary in order to pursue effectively a particular goal and the process that a citizen contends he is due before he is deprived of a constitutional right. . . .

IV

It is beyond question that substantial interests lie on both sides of the scale in this case. Hamdi's "private interest . . . affected by the official action" is the most elemental of liberty interests—the interest in being free from physical detention by one's own government. . . .

Nor is the weight on this side of the . . . scale offset by the circumstances of war or the accusation of treasonous behavior, for "[i]t is clear that commitment for *any* purpose constitutes a significant deprivation of liberty that requires due process protection." . . . Indeed, as *amicus* briefs from media and relief organizations emphasize, the risk of erroneous deprivation of a citizen's liberty in the absence of sufficient process here is very real. . . . Moreover, as critical as the Government's interest may be in detaining those who actually pose an immediate threat to the national security of the United States during ongoing international conflict, history and common sense teach us that an unchecked system of detention carries the potential to become a means for oppression and abuse of others who do not present that sort of threat. . . . We reaffirm today the fundamental nature of a citizen's right to be free from involuntary confinement by his own government without due process of law, and we weigh the opposing governmental interests against the curtailment of liberty that such confinement entails.

On the other side of the scale are the weighty and sensitive government interests in ensuring that those who have in fact fought with the enemy during a war do not return to battle against the United States. As discussed above, the law of war and the realities of combat may render such detentions both necessary and appropriate, and our due process analysis need not blink at those realities. . . .

Striking the proper constitutional balance here is of great importance to the Nation during this period of on-going combat. But it is equally vital that our calculus not give short shrift to the values that this country holds dear or to the privilege that is American citizenship. It is during our most challenging and uncertain moments that our Nation's commitment to due process is most severely tested; and it is in those times that we must preserve our commitment at home to the principles for which we fight abroad. . . .

We therefore hold that a citizen-detainee seeking to challenge his classification as an enemy combatant must receive notice of the factual basis for his classification, and a fair opportunity to rebut the Government's factual assertions before a neutral decisionmaker. . . .

At the same time, the exigencies of the circumstances may demand that, aside from these core elements, enemy combatant proceedings may be tailored to alleviate their uncommon potential to burden the Executive at a time of ongoing military conflict. Hearsay, for example, may need to be accepted as the most reliable available evidence from the Government in such a proceeding. Likewise, the Constitution would not be offended by a presumption in favor of the Government's evidence, so long as that presumption remained a rebuttable one and fair opportunity for rebuttal were provided. . . .

We think it unlikely that this basic process will have the dire impact on the central functions of war-making that the Government forecasts. The parties agree that initial captures on the battlefield need not receive the process we have discussed here; that process is due only when the determination is made to *continue* to hold those who have been seized. . . .

The judgment of the United States Court of Appeals for the Fourth Circuit is vacated, and the case is remanded for further proceedings.

It is so ordered.

JUSTICE SOUTER, with whom JUSTICE GINSBURG joins, concurring in part, dissenting in part, and concurring in the judgment: . . . In these proceedings on Hamdi's petition, he seeks to challenge the facts claimed by the Government as the basis for holding him as an enemy combatant. And in this Court he presses the distinct argument that the Government's claim, even if true, would not implicate any authority for holding him that would satisfy 18 U.S.C. §4001(a) (Non-Detention Act [1971]), which bars imprisonment or detention of a citizen "except pursuant to an Act of Congress."

The Government responds that Hamdi's incommunicado imprisonment as an enemy combatant seized on the field of battle falls within the President's power as Commander in Chief under the laws and usages of war, and is in any event authorized by two statutes. Accordingly, the Government contends that Hamdi has no basis for any challenge by petition for habeas except to his own status as an enemy combatant; and even that challenge may go no further than to enquire whether "some evidence" supports Hamdi's designation. . . .

The plurality rejects any such limit on the exercise of habeas jurisdiction and so far I agree with its opinion. The plurality does, however, accept the Government's position that if Hamdi's designation as an enemy combatant is correct, his detention . . . is authorized by an Act of Congress as required by §4001(a), that is, by the Authorization for Use of Military Force, 115 Stat. 224 (hereinafter Force Resolution). Here, I disagree and respectfully dissent. The Government has failed to demonstrate that the Force Resolution authorizes the detention complained of here even on the facts the Government claims. If the Government raises nothing further than the record now shows, the Non-Detention Act entitles Hamdi to be released.

I

The threshold issue is how broadly or narrowly to read the Non-Detention Act, the tone of which is severe: "No citizen shall be imprisoned or otherwise detained by the United States except pursuant to an Act of Congress." Should the severity of the Act be relieved when the Government's stated factual justification for incommunicado detention is a war on terrorism, so that the Government may be said to act "pursuant" to congressional terms that fall short of explicit authority to imprison individuals? . . . For a number of reasons, the prohibition within §4001(a) has to be read broadly to accord the statute a long reach and to impose a burden of justification on the Government.

First, the circumstances in which the Act was adopted point the way to this interpretation. The provision superseded a cold-war statute, the Emergency Detention Act of 1950 which had authorized the Attorney General, in time of emergency, to detain anyone reasonably thought likely to engage in espionage or sabotage. That statute was repealed in 1971 out of fear that it could authorize a repetition of the World War II internment of citizens of Japanese ancestry. . . .

The fact that Congress intended to guard against a repetition of the World War II internments when it repealed the 1950 statute and gave us §4001(a) provides a powerful reason to think that §4001(a) was meant to require clear congressional authorization before any citizen can be placed in a cell. . . .

Second, when Congress passed §4001(a) it was acting in light of an interpretive regime that subjected enactments limiting liberty in wartime to the requirement of a clear statement and it presumably intended §4001(a) to be read accordingly. This need for clarity was unmistakably expressed in *Ex parte Endo*, decided the same day as *Korematsu*. *Endo* began with a petition for habeas corpus by an interned citizen claiming to be loyal and law-abiding and thus "unlawfully detained." The petitioner was held entitled to habeas relief in an opinion that set out this principle for scrutinizing wartime statutes in derogation of customary liberty:

> "In interpreting a wartime measure we must assume that [its] purpose was to allow for the greatest possible accommodation between . . . liberties and the exigencies of war. We must assume, when asked to find implied powers in a grant of legislative or executive authority, that the law makers intended to place no greater restraint on the citizen than was clearly and unmistakably indicated by the language they used."

Congress's understanding of the need for clear authority before citizens are kept detained is itself

therefore clear, and §4001(a) must be read to have teeth in its demand for congressional authorization.

Finally, even if history had spared us the cautionary example of the internments in World War II, even if there had been no *Korematsu*, and *Endo* had set out no principle of statutory interpretation, there would be a compelling reason to read §4001(a) to demand manifest authority to detain before detention is authorized. The defining character of American constitutional government is its constant tension between security and liberty, serving both by partial helpings of each. In a government of separated powers, deciding finally on what is a reasonable degree of guaranteed liberty whether in peace or war (or some condition in between) is not well entrusted to the Executive Branch of Government, whose particular responsibility is to maintain security. . . .

II

Under this principle of reading §4001(a) robustly to require a clear statement of authorization to detain, none of the Government's arguments suffices to justify Hamdi's detention.

Justice Scalia, with whom Justice Stevens joins, dissenting: Petitioner, a presumed American citizen, has been imprisoned without charge or hearing in the Norfolk and Charleston Naval Brigs for more than two years, on the allegation that he is an enemy combatant who bore arms against his country for the Taliban. His father claims to the contrary, that he is an inexperienced aid worker caught in the wrong place at the wrong time. This case brings into conflict the competing demands of national security and our citizen's constitutional right to personal liberty. Although I share the Court's evident unease as it seeks to reconcile the two, I do not agree with its resolution.

. . . Where the Government accuses a citizen of waging war against it, our constitutional tradition has been to prosecute him in federal court for treason or some other crime. Where the exigencies of war prevent that, the Constitution's Suspension Clause, Art. I §9, cl. 2, allows Congress to relax the usual protections temporarily. Absent suspension, however, the Executive's assertion of military exigency has not been thought sufficient to permit detention without charge. No one contends that the congressional Authorization for Use of Military Force, on which the Government relies to justify its actions here, is an implementation of the Suspension Clause. Accordingly, I would reverse the decision below.

I

The very core of liberty secured by our Anglo-Saxon system of separated powers has been freedom from indefinite imprisonment at the will of the Executive. . . . The gist of the Due Process Clause, as understood at the founding and since, was to force the Government to follow those common-law procedures traditionally deemed necessary before depriving a person of life, liberty, or property. When a citizen was deprived of liberty because of alleged criminal conduct, those procedures typically required committal by a magistrate followed by indictment and trial. . . .

These due process rights have historically been vindicated by the writ of habeas corpus. In England before the founding, the writ developed into a tool for challenging executive confinement. . . .

The writ of habeas corpus was preserved in the Constitution—the only common-law writ to be explicitly mentioned. See Art. I, §9, cl. 2. Hamilton lauded "the establishment of the writ of *habeas corpus*" . . . as a means to protect against "the practice of arbitrary imprisonments . . . in all ages, [one of] the favourite and most formidable instruments of tyranny." . . .

II

The allegations here, of course, are no ordinary accusations of criminal activity. Yaser Esam Hamdi has been imprisoned because the Government believes he participated in the waging of war against the United States. The relevant question, then, is whether there is a different, special procedure for imprisonment of a citizen accused of wrongdoing *by aiding the enemy in wartime.*

A

Justice O'Connor, writing for a plurality of this Court, asserts that captured enemy combatants (other than those suspected of war crimes) have traditionally been detained until the cessation of hostilities and then released. That is probably an accurate description of wartime practice with respect to enemy *aliens.* The tradition with respect to American citizens, however, has been quite different. Citizens aiding the enemy have been treated as traitors subject to the criminal process. . . .

Subjects accused of levying war against the King were routinely prosecuted for treason. The Founders inherited the understanding that a citizen's levying war against the Government was to be punished criminally. The Constitution provides: "Treason against the United States, shall consist only in levying War against them, or in adhering to their Enemies, giving them Aid and Comfort" and establishes a heightened proof requirement (two witnesses) in order to "convic[t]" of that offense. Art. III, §3, cl. 1.

In more recent times, too, citizens have been charged and tried in Article III courts for acts of war against the United States, even when their noncitizen co-conspirators were not. For example, two American citizens alleged to have participated during World War I in a spying conspiracy on behalf of Germany were tried in federal court. . . .

The modern treason statute is 18 U.S.C. §2381; it basically tracks the language of the constitutional provision. Other provisions of Title 18 criminalize various acts of warmaking and adherence to the enemy. . . . The only citizen other that Hamdi known to be imprisoned in connection with military hostilities in Afghanistan against the United States *was* subjected to criminal process and convicted upon a guilty plea.

B

There are times when military exigency renders resort to the traditional criminal process impracticable. English law accommodated such exigencies by allowing legislative suspension of the writ of habeas corpus for brief periods. . . .

Our Federal Constitution contains a provision explicitly permitting suspension, but limiting the situations in which it may be invoked: "The privilege of the Writ of Habeas Corpus shall not be suspended, unless when in Cases of Rebellion or Invasion the public Safety may require it." Art. I, §9, cl. 2. Although this provision does not state that suspension must be effected by, or authorized by, a legislative act, it has been so understood, consistent with English practice and the Clause's placement in Article I. . . .

III

It follows from what I have said that Hamdi is entitled to a habeas decree requiring his release unless (1) criminal proceedings are promptly brought, or (2) Congress has suspended the writ of habeas corpus. A suspension of the writ could, of course, lay down conditions for continued detention, similar to those that today's opinion prescribes under the Due Process Clause. But there is a world of difference between the people's representatives' determining the need for that suspension (and prescribing the conditions for it), and this Court's doing so.

The plurality finds justification for Hamdi's imprisonment in the Authorization for Use of Military Force. . . . This is not remotely a congressional suspension of the writ, and no one claims that it is. Contrary to the plurality's view, I do not think this statute even authorizes detention of a citizen with the clarity necessary to satisfy the interpretive canon that statutes should be construed so as to avoid grave constitutional concerns. . . . But even if it did, I would not permit it to overcome Hamdi's entitlement to habeas corpus relief. The Suspension Clause of the Constitution, which carefully circumscribes the conditions under which the writ can be withheld, would be a sham if it could be evaded by congressional prescription of requirements *other than the common-law requirement of committal for criminal prosecution* that render the writ, though available, unavailing. If the Suspension Clause does not guarantee the citizen that he will either be tried or released, unless the conditions for suspending the writ exist and the grave action of suspending the writ has been taken; if it merely guarantees the

citizen that he will not be detained unless Congress by ordinary legislation says he can be detained; it guarantees him very little indeed.

It should not be thought, however, that the plurality's evisceration of the Suspension Clause augments, principally, the power of Congress. As usual, the major effect of its constitutional improvisation is to increase the power of the Court. Having found a congressional authorization for detention of citizens where none clearly exists; and having discarded the categorical procedural protection of the suspension Clause; the plurality then proceeds, under the guise of the Due Process Clause, to prescribe what procedural protections *it* thinks appropriate. It "weigh[s] the private interest . . . against the Government's asserted interest," and—just as though writing a new Constitution—comes up with an unheard-of system in which the citizen rather than the Government bears the burden of proof, testimony is by hearsay rather than live witnesses, and the presiding officer may well be a "neutral" military officer rather than judge and jury. . . .

Having distorted the Suspension Clause, the plurality finishes up by transmogrifying the Great Writ—disposing of the present habeas petition by remanding for the District Court to "engag[e] in a factfinding process that is both prudent and incremental." . . . This judicial remediation of executive default is unheard of. The role of habeas corpus is to determine the legality of executive detention, not to supply the omitted process necessary to make it legal. . . . It is not the habeas court's function to make illegal detention legal by supplying a process that the Government could have provided, but chose not to. If Hamdi is being imprisoned in violation of the Constitution (because without due process of law), then his habeas petition should be granted; the Executive may than hand him over to the criminal authorities, whose detention for the purpose of prosecution will be lawful, or else must release him. . . .

IV

Many think it not only inevitable but entirely proper that liberty give way to security in times of national crisis—that, at the extremes of military exigency, *inter arma silent leges* [war silences law]. Whatever the general merits of the view that war silences law or modulates its voice, that view has no place in the interpretation and application of a Constitution designed precisely to confront war and, in a manner that accords with democratic principles, to accommodate it. Because the Court has proceeded to meet the current emergency in a manner the Constitution does not envision, I respectfully dissent.

Justice Thomas, dissenting: The Executive Branch, acting pursuant to the powers vested in the President by the Constitution and with explicit congressional approval, has determined that Yaser Hamdi is an enemy combatant and should be detained. This detention falls squarely within the Federal Government's war powers, and we lack the expertise and capacity to second-guess that decision. As such, petitioner's habeas challenge should fail, and there is no reason to remand the case. The plurality reaches a contrary conclusion by failing adequately to consider basic principles of the constitutional structure as it relates to national security and foreign affairs and by using [a] balancing scheme. . . . I do not think that the Federal Government's war powers can be balanced away by this Court. Arguably, Congress could provide for additional procedural protections, but until it does, we have no right to insist upon them. But even if I were to agree with the general approach the plurality takes, I could not accept the particulars. The plurality utterly fails to account for the Government's compelling interests and for our own institutional inability to weigh competing concerns correctly. I respectfully dissent.

I

. . . The Founders intended that the President have primary responsibility—along with the necessary power—to protect the national security and to conduct the Nation's foreign relations. They did so principally because the structural advantages of a unitary Executive are essential in these domains. . . . This Court has long recognized these features and has accordingly held that the President has *constitutional* authority to protect the national security and that this authority carries with it broad discretion. . . .

Congress, to be sure, has a substantial and essential role in both foreign affairs and national security. But it is crucial to recognize that *judicial* interference in these domains destroys the purpose of vesting primary responsibility in a unitary Executive. . . . First, with respect to certain decisions relating to national security and foreign affairs, the courts simply lack the relevant information and expertise to second-guess determinations made by the President based on information properly withheld. Second, even if the courts could compel the Executive to produce the necessary information, such decisions are simply not amenable to judicial determination because "[t]hey are delicate, complex, and involve large elements of prophecy." Third, the Court . . . has correctly recognized the primacy of the political branches in the foreign-affairs and national-security contexts. . . .

Finally, and again for the same reasons, where "the President acts pursuant to an express or implied authorization from Congress, he exercises not only his powers but also those delegated by Congress, [and] in such a case the executive action 'would be supported by the strongest of presumptions and the widest latitude of judicial interpretation.' " *Dames & Moore*, at 668. . . . This deference extends to the President's determination of all the factual predicated necessary to conclude that a given action is appropriate. . . .

I acknowledge that the question whether Hamdi's executive detention is lawful is a question properly resolved by the Judicial Branch, though the question comes to the Court with the strongest presumptions in favor of the Government. The plurality agrees that Hamdi's detention is lawful if he is an enemy combatant. But the question whether Hamdi is actually an enemy combatant is "of a kind for which the Judiciary has neither aptitude, facilities nor responsibility and which has long been held to belong in the domain of political power not subject to judicial intrusion or inquiry." *Chicago & Southern Air Lines*, 333 U.S., at 111. That is, although it is appropriate for the Court to determine the judicial question whether the President has the asserted authority, we lack the information and expertise to question whether Hamdi is actually an enemy combatant, a question the resolution of which is committed to other branches. . . .

For these reasons, I would affirm the judgment of the Court of Appeals.

REVIEW AND DISCUSSION QUESTIONS

1. What are the facts of this case? How did the case unfold at the lower judicial levels before it reached the Supreme Court? What do you see as the most important legal issues in the case?
2. Justice O'Connor and the plurality contend that Congress authorized the detention of citizens as enemy combatants when it passed the AUMF. Justice Souter disagrees, relying on a strict application of the 1971 Non-Detention Act. With whom do you side, and why?
3. The plurality argues that the Mobbs declaration is an insufficient basis for holding Hamdi and that habeas corpus entails that he be granted a meaningful opportunity to contest his detention. Justice Thomas believes that the Court should defer to the judgment of the executive branch in this matter. With whom do you side, and why? Is there a danger that upholding Hamdi's rights will damage the government's ability to wage war effectively?
4. Explain and critically assess Justice Scalia's argument the Hamdi must either be charged with a criminal offense or released.
5. The plurality says that although Hamdi should have an opportunity to rebut his classification as an enemy combatant, various procedural rules may be relaxed because of the special, military circumstances. What modifications to ordinary criminal procedure does the plurality envision? Why do you think it takes this position? Explain why Scalia criticizes the plurality on this matter.
6. Explain what you think habeas corpus, due process, and the rule of law require in this case. Which of the four justices—O'Connor, Souter, Scalia, or Thomas—presents the most compelling legal argument?
7. Hamdi is an American citizen. What rights, if any, should non-citizens who are detained as enemy combatants have? Should they be granted a meaningful opportunity to contest their detention before a neutral decision maker? Explain why or why not.

CHAPTER 3

The Moral Force of Law

Essays in this section focus on two interrelated topics: legal obligation and civil disobedience. Is there reason to obey the law, merely because it is the law? To introduce this topic, we chose what is probably the best known discussion of legal obligation in the history of Western philosophy: Plato's *Crito.* The two subsequent essays explore the related topic of civil disobedience: What is it? When is it justified? Should officials charged with enforcing the law always prosecute civil disobedients? Would not prosecuting them be unfair to those who do obey the law? Embedded in these discussions are other topics that will emerge in later sections, including the nature of law itself and the connections between a law's being unjust and its status as valid law.

Crito

Plato

Plato was born into an aristocratic family in Athens around 427 B.C. Socrates (470–399 B.C.) was a friend of Plato's family from the time Plato was a schoolboy. Although a young Athenian of Plato's class would normally have pursued a political career, Plato chose philosophy instead and achieved fame as Socrates' most talented student and one of the world's greatest philosophers.

In 387 B.C., Plato founded the Academy, a school of higher education and research, which existed for more than 900 years, and he continued as its head until his death in 347 B.C. at the age of eighty. Plato wrote some twenty-five dialogues, including "Crito" and the book-length "Republic."

"Crito" begins after Socrates has been condemned to death for corrupting the youth through his teaching. Socrates' friend Crito tries to convince Socrates to allow him to save his life by breaking him out of jail, but Socrates will agree only if he can be convinced that violating the law would be just. "Crito" thus sets the stage for centuries of debate over the nature and extent of a citizen's legal obligation. (Socrates, in fact, refused to disobey the law and was executed.)

SCENE: *THE PRISON OF SOCRATES*

SOCRATES: Dear Crito, . . . I am and always have been one of those natures who must be guided by reason, whatever the reason may be which upon reflection appears to me to be the best; and now that this fortune has come upon me, I can not put away the reasons which I have before given: the principles which I have hitherto honored and revered I still honor, and unless we can find other and better principles on the instant, I am certain not to agree with you; no, not even if the power of the multitude could inflict many more imprisonments, confiscations, deaths, frightening us like children with hobgoblin terrors. . . . Now you, Crito, are a disinterested person who are not going to die tomorrow—at least, there is no human probability of this, and you are therefore not liable to be deceived by the circumstances in which you are placed. Tell me then, whether I am right in saying that some opinions, and the opinions of some men only, are to be valued, and other opinions and the opinions of other men, are not to be valued. I ask you whether I was right in maintaining this?

CRITO: Certainly.

SOCRATES: Very good; and is not this true, Crito, of other things which we need not separately enumerate? In the matter of just and unjust, fair and foul, good and evil, which are the subjects of our present consultation, ought we to follow the opinion of the many and to fear them; or the opinion of the one man who has understanding, and whom we ought to fear and reverence more than all the rest of the world; and whom deserting we shall destroy and injure that principle in us which may be assumed to be improved by justice and deteriorated by injustice;—is there not such a principle?

CRITO: Certainly there is, Socrates.

SOCRATES: Take a parallel instance:—if, acting under the advice of men who have no understanding, we destroy that which is improvable by health and deteriorated by disease—when that has been destroyed, I say, would life be worth having? And that is—the body?

CRITO: Yes.

SOCRATES: Could we live, having an evil and corrupted body?

CRITO: Certainly not.

SOCRATES: And will life be worth having, if that higher part of man be depraved, which is improved by justice and deteriorated by injustice? Do we suppose that principle, whatever it may be in man, which has to do with justice and injustice, to be inferior to the body?

CRITO: Certainly not.

SOCRATES: More honored, then?

CRITO: Far more honored.

SOCRATES: Then, my friend, we must not regard what the many say of us; but what he, the one man who has understanding of just and unjust, will say, and what the truth will say. And therefore you begin in error when you suggest that we should regard the opinion of the many about just and unjust, good and evil, honorable and dishonorable.—Well, some one will say, "but the many can kill us."

CRITO: Yes, Socrates; that will clearly be the answer.

SOCRATES: That is true: but still I find with surprise that the old argument is, as I conceive, unshaken as ever. And I should like to know whether I may say the same of another proposition—that not life, but a good life, is to be chiefly valued?

CRITO: Yes, that also remains.

SOCRATES: And a good life is equivalent to a just and honorable one—that holds also?

CRITO: Yes, that holds.

SOCRATES: From these premises I proceed to argue the question whether I ought not to try and escape without the consent of the Athenians: and if I am clearly right in escaping, then I will make the attempt; but if not, I will abstain. The other considerations which you mention, of money and loss of character and the duty of educating children, are, as I fear, only the doctrines of the multitude, who would be as ready to call people to life, if they were able, as they are to put them to death—and with as little reason. But now, since the argument has thus far prevailed, the only question which remains to be considered is, whether we shall do rightly either in escaping or in suffering others to aid in our escape and paying them in money and thanks, or whether we shall not do rightly; and if the latter, then death or any other calamity which may ensue on my remaining here must not be allowed to enter into the calculation.

Crito: I think that you are right, Socrates; how then shall we proceed?

Socrates: Let us consider the matter together, and do you either refute me if you can, and I will be convinced; or else cease, my dear friend, from repeating to me that I ought to escape against the wishes of the Athenians: for I am extremely desirous to be persuaded by you, but not against my own better judgment. And now please to consider my first position, and do your best to answer me.

Crito: I will do my best.

Socrates: Are we to say that we are never intentionally to do wrong, or that in one way we ought and in another way we ought not to do wrong, or is doing wrong always evil and dishonorable, as I was just now saying, and as has been already acknowledged by us? Are all our former admissions which were made within a few days to be thrown away? And have we, at our age, been earnestly discoursing with one another all our life long only to discover that we are no better than children? Or are we to rest assured, in spite of the opinion of the many, and in spite of consequences whether better or worse, of the truth of what was then said, that injustice is always an evil and dishonor to him who acts unjustly? Shall we affirm that?

Crito: Yes.

Socrates: Then we must do no wrong?

Crito: Certainly not.

Socrates: Nor when injured injure in return, as the many imagine; for we must injure no one at all?

Crito: Clearly not.

Socrates: Again, Crito, may we do evil?

Crito: Surely not, Socrates.

Socrates: And what of doing evil in return for evil, which is the morality of the many—is that just or not?

Crito: Not just.

Socrates: For doing evil to another is the same as injuring him?

Crito: Very true.

Socrates: Then we ought not to retaliate or render evil for evil to any one, whatever evil we may have suffered from him. But I would have you consider, Crito, whether you really mean what you are saying. For this opinion has never been held, and never will be held, by any considerable number of persons; and those who are agreed and those who are not agreed upon this point have no common ground, and can only despise one another when they see how widely they differ. Tell me, then, whether you agree with and assent to my first principle, that neither injury nor retaliation nor warding off evil by evil is ever right. And shall that be the premise of our argument? Or do you decline and dissent from this? For this has been of old and is still my opinion; but, if you are of another opinion, let me hear what you have to say. If, however, you remain of the same mind as formerly, I will proceed to the next step.

Crito: You may proceed, for I have not changed my mind.

Socrates: Then I will proceed to the next step, which may be put in the form of a question:—Ought a man to do what he admits to be right, or ought he to betray the right?

Crito: He ought to do what he thinks right.

Socrates: But if this is true, what is the application? In leaving the prison against the will of the Athenians, do I wrong any? or rather do I not wrong those whom I ought least to wrong? Do I not desert the principles which were acknowledged by us to be just? What do you say?

Crito: I can not tell, Socrates, for I do not know.

Socrates: Then consider the matter in this way:—Imagine that I am about to play truant (you may call the proceeding by any name which you like), and the laws and the government come and interrogate me: "Tell us, Socrates," they say; "what are you about? are you going by an act of yours to overturn us—the laws and the whole state, as far as in you lies? Do you imagine that a state can subsist and not be overthrown, in which the decisions of law have no power, but are set aside and overthrown by individuals?" What will be our answer, Crito, to these and the like words? Any one, and especially a clever rhetorician, will have a good deal to urge about the evil of setting aside the law which requires a sentence to be carried out; and we might reply, "Yes; but the state has injured us and given an unjust sentence." Suppose I say that?

Crito: Very good, Socrates.

Socrates: "And was that our agreement with you?" the law would say; "or were you to abide by the sentence of the state?" And if I were to express astonishment at their saying this, the law would probably add: "Answer, Socrates, instead of opening your eyes: you are in the habit of asking

and answering questions. Tell us what complaint you have to make against us which justifies you in attempting to destroy us and the state? In the first place did we not bring you into existence? Your father married your mother by our aid and begat you. Say whether you have any objection to urge against those of us who regulate marriage?" None, I should reply. "Or against those of us who regulate the system of nurture and education of children in which you were trained? Were not the laws, who have the charge of this, right in commanding your father to train you in music and gymnastic?" Right, I should reply, "Well then, since you were brought into the world and nurtured and educated by us, can you deny in the first place that you are our child and slave, as your fathers were before you? And if this is true you are not on equal terms with us; nor can you think that you have a right to do to us what we are doing to you. Would you have any right to strike or revile or do any other evil to a father or to your master, if you had one, when you have been struck or reviled by him, or received some other evil at his hands?—you would not say this? And because we think right to destroy you, do you think that you have any right to destroy us in return, and your country as far as in you lies? And will you, O professor of true virtue, say that you are justified in this? Has a philosopher like you failed to discover that our country is more to be valued and higher and holier far than mother or father or any ancestor, and more to be regarded in the eyes of the gods and of men of understanding? Also to be soothed, and gently and reverently entreated when angry, even more than a father, and if not persuaded, obeyed? And when we are punished by her, whether with imprisonment or stripes, the punishment is to be endured in silence; and if she lead us to wounds or death in battle, thither we follow as is right; neither may any one yield or retreat or leave his rank, but whether in battle or in a court of law, or in any other place, he must do what his city and his country order him; or he must change their view of what is just; and if he may do no violence to his father or mother, much less may he do violence to his country." What answer shall we make to his, Crito? Do the laws speak truly, or do they not?

Crito: I think that they do.

Socrates: Then the laws will say: "Consider, Socrates, if this is true, that in your present attempt you are going to do us wrong. For, after having brought you into the world, and nurtured and educated you, and given you and every other citizen a share in every good that we had to give, we further proclaim and give the right to every Athenian, that if he does not like us when he has come of age and has seen the ways of the city, and made our acquaintance, he may go where he pleases and take his goods with him; and none of us laws will forbid him or interfere with him. Any of you who does not like us and the city, and who wants to go to a colony or to any other city, may go where he likes, and take his goods with him. But he who has experience of the manner in which we order justice and administer the state, and still remains, has entered into an implied contract that he will do as we command him. And he who disobeys us is, as we maintain, thrice wrong; first, because in disobeying us he is disobeying his parents; secondly, because we are the authors of his education; thirdly, because he has made an agreement with us that he will duly obey our commands; and he neither obeys them nor convinces us that our commands are wrong; and we do not rudely impose them, but give them the alternative of obeying or convincing us;—that is what we offer, and he does neither. These are the sort of accusations to which, as we were saying, you, Socrates, will be exposed if you accomplish your intentions; you, above all other Athenians." Suppose I ask, why is this? they will justly retort upon me that I above all other men have acknowledged the agreement. "There is clear proof," they will say, "Socrates, that we and the city were not displeasing to you. Of all Athenians you have been the most constant resident in the city, which, as you never leave, you may be supposed to love. For you never went out of the city either to see the games, except once when you went to the Isthmus, or to any other place unless you were on military service; nor did you travel as other men do. Nor had you any curiosity to know other states or their laws: your affections did not go beyond us and our state; we were your special favorites, and you acquiesced in our government of you; and this is the state in which you begat your children, which is a proof of your satisfaction. Moreover, you might, if you had liked, have fixed the penalty at banishment in the

course of the trial—the state which refuses to let you go now would have let you go then. But you pretended that you preferred death to exile, and that you were not grieved at death. And now you have forgotten these fine sentiments, and pay no respect to us the laws, of whom you are the destroyer; and are doing what only a miserable slave would do, running away and turning your back upon the compacts and agreements which you made as a citizen. And first of all answer this very question: Are we right in saying that you agreed to be governed according to us in deed, and not in word only? Is that true or not?" How shall we answer that, Crito? Must we not agree?

Crito: There is no help, Socrates.

Socrates: Then will they not say: "You, Socrates, are breaking the covenants and agreements which you made with us at your leisure, not in any haste or under any compulsion or deception, but having had seventy years to think of them, during which time you were at liberty to leave the city, if we were not to your mind, or if our covenants appeared to you to be unfair. You had your choice, and might have gone either to Lacedaemon or Crete, which you often praise for their good government, or to some other Hellenic or foreign state. Whereas you, above all other Athenians, seemed to be so fond of the state, or, in other words, of us her laws (for who would like a state that has no laws), that you never stirred out of her; the halt, the blind, the maimed were not more stationary in her than you were. And now you run away and forsake your agreements. Not so, Socrates, if you will take our advice; do not make yourself ridiculous by escaping out of the city.

"For just consider, if you transgress and err in this sort of way, what good will you do either to yourself or to your friends? That your friends will be driven into exile and deprived of citizenship, or will lose their property, is tolerably certain; and you yourself, if you fly to one of the neighboring cities, as, for example, Thebes or Megara, both of which are well-governed cities, will come to them as an enemy, Socrates, and their government will be against you, and all patriotic citizens will cast an evil eye upon you as a subverter of the laws, and you will confirm in the minds of the judges the justice of their own condemnation of you. For he who is a corruptor of the laws is more than likely to be corruptor of the young and foolish portion of mankind. Will you then flee from well-ordered cities and virtuous men? And is existence worth having on these terms? Or will you go to them without shame, and talk to them, Socrates? And what will you say to them? What you say here about virtue and justice and institutions and laws being the best things among men. Would that be decent of you? Surely not. But if you go away from well-governed states to Crito's friends in Thessaly, where there is a great disorder and license, they will be charmed to have the tale of your escape from prison, set off with ludicrous particulars of the manner in which you were wrapped in a goatskin or some other disguise, and metamorphosed as the fashion of runaways is—that is very likely; but will there be no one to remind you that in your old age you violated the most sacred laws from a miserable desire of a little more life. Perhaps not, if you keep them in a good temper; but if they are out of temper you will hear many degrading things; you will live, but how?—as the flatterer of all men, and the servant of all men; and doing what?—eating and drinking in Thessaly, having gone abroad in order that you may get a dinner. And where will be your fine sentiments about justice and virtue then? Say that you wish to live for the sake of your children, that you may bring them up and educate them—will you take them into Thessaly and deprive them of Athenian citizenship? Is that the benefit which you would confer upon them? Or are you under the impression that they will be better cared for and educated here if you are still alive, although absent from them; for that your friends will take care of them? Do you fancy that if you are an inhabitant of Thessaly they will take care of them, and if you are an inhabitant of the other world they will not take care of them? Nay; but if they who call themselves friends are truly friends, they surely will.

"Listen, then, Socrates, to us who have brought you up. Think not of life and children first, and of justice afterwards, but of justice first, that you may be justified before the princes of the world below. For neither will you nor any that belong to you be happier or holier or juster in this life, or happier in another, if you do as Crito bids. Now you depart in innocence, a sufferer and not a doer of evil; a victim, not of the laws, but of men. But if you go forth, returning evil for evil, and injury for injury, breaking the covenants and agreements which you have made with us, and wronging those whom you ought least to wrong, that is to say, yourself, your

friends, your country, and us, we shall be angry with you while you live, and our brethren, the laws in the world below, will receive you as an enemy; for they will know that you have done your best to destroy us. Listen, then, to us and not to Crito."

This is the voice which I seem to hear murmuring in my ears, like the sound of the flute in the ears of the mystic; that voice, I say, is humming in my ears, and prevents me from hearing any other. And I know that anything more which you may say will be vain. Yet speak, if you have anything to say.

CRITO: I have nothing to say, Socrates.

SOCRATES: Then let me follow the intimations of the will of God.

REVIEW AND DISCUSSION QUESTIONS

1. Elaborate on Socrates' notion that the law is like a parent and so deserves obedience from its citizens.
2. Sometimes we can lead others to expect us to behave in a certain way even though we never overtly promise to do so. Give an example of such behavior, and then discuss whether Socrates has done anything that can fairly be interpreted as leading others to expect him to obey the law.
3. At his trial, Socrates made an argument in his own defense. Is that fact relevant to the present issue?
4. Obviously, Socrates has benefited from living in Athens—benefits that he gladly accepted. Does that constitute an "implied contract" to obey?
5. What is gratitude? Does Socrates think part of his obligation to remain is based on gratitude? Explain.

The Justification of Civil Disobedience

John Rawls

Ours is a multicultural, diverse society that includes many incompatible visions of the good life, of religion, and of the meaning and purposes of human existence. We also have a tradition of respecting individual rights, as well as a commitment to majority rule. Given that diversity and those commitments, it seems inevitable that there will be conflicts over the nature and extent of rights as well as about the limits on governmental power. Such conflicts can lead to disobedience, and thus to philosophical questions about the basis of legal obligation, the justification of civil disobedience, and, ultimately, the nature of social justice itself.

Few twentieth-century philosophers have had as much influence on our thinking about these issues as John Rawls. His seminal work, *A Theory of Justice,* offers a new and important work in the social contract tradition. Rawls argues that principles governing a constitution and laws enacted under it should be chosen behind a "veil of ignorance" in which people ignore their social class, race, and religion, as well as any other facts that are morally arbitrary and that, if known, would allow the laws to be tailored so as to advantage one group over another. He calls the theory "justice as fairness" because the veil of ignorance (corresponding to the traditional social contract) assures that the chosen principles are fair. Forced to choose a constitution and laws behind such a veil of ignorance, Rawls argues that people would secure for themselves basic rights, assure equality of opportunity, and allow economic inequality only if everybody benefits from it, including the least advantaged.

In the following essay, John Rawls discusses the related questions of legal obligation and civil disobedience within the context of social contract theory. Even assuming the constitution is just, there is no assurance laws enacted in accordance with its procedures will also be just. Nonetheless, Rawls argues, citizens may have an obligation to obey that is based on their having benefited from the (generally) just regime. He concludes with a discussion of the role of civil disobedience in a constitutional democracy, including the circumstances in which it would be justified. John Rawls was professor of philosophy at Harvard University.

INTRODUCTION

I should like to discuss briefly, and in an informal way, the grounds of civil disobedience in a constitutional democracy. Thus, I shall limit my remarks to the conditions under which we may, by civil disobedience, properly oppose legally established democratic authority; I am not concerned with the situation under other kinds of government nor, except incidentally, with other forms of resistance. My thought is that in a reasonably just (though of course not perfectly just) democratic regime, civil disobedience, when it is justified, is normally to be understood as a political action which addresses the sense of justice of the majority in order to urge reconsideration of the measures protested and to warn that in the firm opinion of the dissenters the conditions of social cooperation are not being honored. This characterization of civil disobedience is intended to apply to dissent on fundamental questions of internal policy, a limitation which I shall follow to simplify our question.

THE SOCIAL CONTRACT DOCTRINE

It is obvious that the justification of civil disobedience depends upon the theory of political obligation in general, and so we may appropriately begin with a few comments on this question. The two chief virtues of social institutions are justice and efficiency, where by the efficiency of institutions I understand their effectiveness for certain social conditions and ends the fulfillment of which is to everyone's advantage. We should comply with and do our part in just and efficient social arrangements for at least two reasons: first of all, we have a natural duty not to oppose the establishment of just and efficient institutions (when they do not yet exist) and to uphold and comply with them (when they do exist); and second, assuming that we have knowingly accepted the benefits of these institutions and plan to continue to do so, and that we have encouraged and expect others to do their part, we also have an obligation to do our share when, as the arrangement requires, it comes our turn. Thus, we often have both a natural duty as well as an obligation to support just and efficient institutions, the obligation arising from our voluntary acts while the duty does not.

Now all this is perhaps obvious enough, but it does not take us very far. Any more particular conclusions depend upon the conception of justice which is the basis of a theory of political obligation. I believe that the appropriate conception, at least for an account of political obligation in a constitutional democracy, is that of the social contract theory from which so much of our political thought derives. If we are careful to interpret it in a suitably general way, I hold that this doctrine provides a satisfactory basis for political theory, indeed even for ethical theory itself, but this is beyond our present concern. The interpretation I suggest is the following: that the principles to which social arrangements must conform, and in particular the principles of justice, are those which free and rational men would agree to in an original position of equal liberty; and similarly, the principles which govern men's relations to institutions and define their natural duties and obligations are the principles to which they would consent when so situated. It should be noted straightway that in this interpretation of the contract theory the principles of justice are understood as the outcome of a hypothetical agreement. They are principles which would be agreed to if the situation of the original position were to arise. There is no mention of an actual agreement nor need such an agreement ever be made. Social arrangements are just or unjust according to whether they accord with the principles for assigning and securing fundamental rights and liberties which would be chosen in the original position. This position is, to be sure, the analytic analogue of the traditional notion of the state of nature, but it must not be mistaken for a historical occasion. Rather it is a hypothetical situation which embodies the basic ideas of the contract doctrine; the description of this situation enables us to work out which principles would be adopted. I must now say something about these matters.

The original doctrine has always supposed that the persons in the original position have equal powers and rights, that is, that they are symmetrically situated with respect to any arrangements for reaching agreement, and that coalitions and the like are excluded. But it is an essential element (which has not been sufficiently observed although it is implicit in Kant's version of the theory) that there are very strong restrictions on what the contracting parties are presumed to

know. In particular, I interpret the theory to hold that the parties do not know their position in society, past, present, or future; nor do they know which institutions exist. Again, they do not know their own place in the distribution of natural talents and abilities, whether they are intelligent or strong, man or woman, and so on. Finally, they do not know their own particular interests and preferences or the system of ends which they wish to advance: they do not know their conception of the good. In all these respects the parties are confronted with a veil of ignorance which prevents any one from being able to take advantage of his good fortune or particular interests or from being disadvantaged by them. What the parties do know (or assume) is that Hume's circumstances of justice obtain: namely, that the bounty of nature is not so generous as to render cooperative schemes superfluous nor so harsh as to make them impossible. Moreover, they assume that the extent of their altruism is limited and that, in general, they do not take an interest in one another's interests. Thus, given the special features of the original position, each man tries to do the best he can for himself by insisting on principles calculated to protect and advance his system of ends whatever it turns out to be.

I believe that as a consequence of the peculiar nature of the original position there would be an agreement on the following two principles for assigning rights and duties and for regulating distributive shares as these are determined by the fundamental institutions of society: first, each person is to have an equal right to the most extensive liberty compatible with a like liberty for all; second, social and economic inequalities (as defined by the institutional structure or fostered by it) are to be arranged so that they are both to everyone's advantage and attached to positions and offices open to all. In view of the content of these two principles and their application of the main institutions of society, and therefore to the social system as a whole, we may regard them as the two principles of justice. Basic social arrangements are just insofar as they conform to these principles, and we can, if we like, discuss questions of justice directly by reference to them. But a deeper understanding of the justification of civil disobedience requires, I think, an account of the derivation of these principles provided by the doctrine of the social contract. Part of our task is to show why this is so.

THE GROUNDS OF COMPLIANCE WITH AN UNJUST LAW

If we assume that in the original position men would agree both to the principle of doing their part when they have accepted and plan to continue to accept the benefits of just institutions (the principle of fairness), and also to the principle of not preventing the establishment of just institutions and of upholding and complying with them when they do exist, then the contract doctrine easily accounts for our having to conform to just institutions.

[*Editors' Note: Rawls has argued elsewhere that legal obligation rests on the duty of "fair play." Though not absolute, the obligation exists whenever there is a "mutually beneficial and just scheme of cooperation" in which the advantages depend on the willing cooperation of nearly everyone. According to Rawls, because social cooperation requires sacrifice on the part of everybody, including a restriction on personal liberty, and because the benefits are "free" in the sense that each person can benefit individually as long as others do their part whether or not he does his part, each citizen is therefore bound to do his fair share. In a constitutional democracy, that means obeying the law.*]

But how does it account for the fact that we are normally required to comply with unjust laws as well? The injustice of a law is not a sufficient ground for not complying with it any more than the legal validity of legislation is always sufficient to require obedience to it. Sometimes one hears these extremes asserted, but I think that we need not take them seriously.

An answer to our question can be given by elaborating the social contract theory in the following way. I interpret it to hold that one is to envisage a series of agreements as follows: first, men are to agree upon the principles of justice in the original position. Then they are to move to a constitutional convention in which they choose a constitution that satisfies the principles of justice already chosen. Finally they assume the role of a legislative body and guided by the principles of justice enact laws

subject to the constraints and procedures of the just constitution. The decisions reached in any stage are binding in all subsequent stages. Now whereas in the original position the contracting parties have no knowledge of their society or of their own position in it, in both a constitutional convention and a legislature, they do know certain general facts about their institutions, for example, the statistics regarding employment and output required for fiscal and economic policy. But no one knows particular facts about his own social class or his place in the distribution of natural assets. On each occasion the contracting parties have the knowledge required to make their agreement rational from the appropriate point of view, but not so much as to make them prejudiced. They are unable to tailor principles and legislation to take advantage of their social or natural position; a veil of ignorance prevents their knowing what this position is. With this series of agreements in mind, we can characterize just laws and policies as those which would be enacted were this whole process correctly carried out.

In choosing a constitution the aim is to find among the just constitutions the one which is most likely, given the general facts about the society in question, to lead to just and effective legislation. The principles of justice provide a criterion for the laws desired; the problem is to find a set of political procedures that will give this outcome. I shall assume that, at least under the normal conditions of a modern state, the best constitution is some form of democratic regime affirming equal political liberty and using some sort of majority (or other plurality) rule. Thus it follows that on the contract theory a constitutional democracy of some sort is required by the principles of justice. At the same time it is essential to observe that the constitutional process is always a case of what we may call imperfect procedural justice: that is, there is no feasible political procedure which guarantees that the enacted legislation is just even though we have (let us suppose) a standard for just legislation. In simple cases, such as games of fair division, there are procedures which always lead to the right outcome (assume that equal shares is fair and let the man who cuts the cake take the last piece). These situations are those of perfect procedural justice. In other cases it does not matter what the outcome is as long as the fair procedure is followed: fairness of the process is transferred to the result (fair gambling is an instance of this). These situations are those of pure procedural justice. The constitutional process, like a criminal trial, resembles neither of these; the result matters and we have a standard for it. The difficulty is that we cannot frame a procedure which guarantees that only just and effective legislation is enacted. Thus even under a just constitution unjust laws may be passed and unjust policies enforced. Some form of the majority principle is necessary but the majority may be mistaken, more or less willfully, in what it legislates. In agreeing to a democratic constitution (as an instance of imperfect procedural justice) one accepts at the same time the principle of majority rule. Assuming that the constitution is just and that we have accepted and plan to continue to accept its benefits, we then have both an obligation and a natural duty (and in any case the duty) to comply with what the majority enacts even though it may be unjust. In this way we become bound to follow unjust laws, not always, of course, but provided the injustice does not exceed certain limits. We recognize that we must run the risk of suffering from the defects of one another's sense of justice; this burden we are prepared to carry as long as it is more or less evenly distributed or does not weigh too heavily. Justice binds us to a just constitution and to the unjust laws which may be enacted under it in precisely the same way that it binds us to any other social arrangement. Once we take the sequence of stages into account, there is nothing unusual in our being required to comply with unjust laws.

It should be observed that the majority principle has a secondary place as a rule of procedure which is perhaps the most efficient one under usual circumstances for working a democratic constitution. The basis for it rests essentially upon the principles of justice and therefore we may, when conditions allow, appeal to these principles against unjust legislation. The justice of the constitution does not insure the justice of laws enacted under it; and while we often have both an obligation and a duty to comply with what the majority legislates (as long as it does not exceed certain limits), there is, of course, no corresponding obligation or duty to regard what the majority enacts as itself just. The right to make law does not guarantee that the

decision is rightly made; and while the citizen submits in his conduct to the judgment of democratic authority, he does not submit his judgment to it. And if in his judgment the enactments of the majority exceed certain bounds of injustice, the citizen may consider civil disobedience. For we are not required to accept the majority's acts unconditionally and to acquiesce in the denial of our and others' liberties; rather we submit our conduct to democratic authority to the extent necessary to share the burden of working a constitutional regime, distorted as it must inevitably be by men's lack of wisdom and the defects of their sense of justice.

THE PLACE OF CIVIL DISOBEDIENCE IN A CONSTITUTIONAL DEMOCRACY

We are now in a position to say a few things about civil disobedience. I shall understand it to be a public, nonviolent, and conscientious act contrary to law usually done with the intent to bring about a change in the policies or laws of the government. Civil disobedience is a political act in the sense that it is an act justified by moral principles which define a conception of civil society and the public good. It rests, then, on political conviction as opposed to a search for self or group interest; and in the case of a constitutional democracy, we may assume that this conviction involves the conception of justice (say that expressed by the contract doctrine) which underlies the constitution itself. That is, in a viable democratic regime there is a common conception of justice by reference to which its citizens regulate their political affairs and interpret the constitution. Civil disobedience is a public art which the dissenter believes to be justified by this conception of justice and for this reason it may be understood as addressing the sense of justice of the majority in order to urge reconsideration of the measures protested and to warn that, in the sincere opinion of the dissenters, the conditions of social cooperation are not being honored. For the principles of justice express precisely such conditions, and their persistent and deliberate violation in regard to basic liberties over any extended period of time cuts the ties of community and invites either submission or forceful resistance. By engaging in civil disobedience a minority leads the majority to consider whether it wants to have its acts taken in this way, or whether, in view of the common sense of justice, it wishes to acknowledge the claims of the minority.

Civil disobedience is also civil in another sense. Not only is it the outcome of a sincere conviction based on principles which regulate civic life, but it is public and nonviolent, that is, it is done in a situation where arrest and punishment are expected and accepted without resistance. In this way it manifests a respect for legal procedures. Civil disobedience expresses disobedience to law within the limits of fidelity to law, and this feature of it helps to establish in the eyes of the majority that it is indeed conscientious and sincere, that it really is meant to address their sense of justice. Being completely open about one's acts and being willing to accept the legal consequences of one's conduct is a bond given to make good one's sincerity, for that one's deeds are conscientious is not easy to demonstrate to another or even before oneself. No doubt it is possible to imagine a legal system in which conscientious belief that the law is unjust is accepted as a defense for noncompliance, and men of great honesty who are confident in one another might make such a system work. But as things are such a scheme would be unstable; we must pay a price in order to establish that we believe our actions have a moral basis in the convictions of the community.

The nonviolent nature of civil disobedience refers to the fact that it is intended to address the sense of justice of the majority and as such it is a form of speech, an expression of conviction. To engage in violent acts likely to injure and to hurt is incompatible with civil disobedience as a mode of address. Indeed, an interference with the basic rights of others tends to obscure the civilly disobedient quality of one's act. Civil disobedience is nonviolent in the further sense that the legal penalty for one's action is accepted and that resistance is not (at least for the moment) contemplated. Nonviolence in this sense is to be distinguished from nonviolence as a religious or pacifist principle. While those engaging in civil disobedience have often held some such principle, there is no necessary connection between it and civil disobedience. For on the interpretation suggested, civil disobedience in a democratic society is best understood as an appeal to the principles of justice, the fundamental conditions of

willing social cooperation among free men, which in the view of the community as a whole are expressed in the constitution and guide its interpretation. Being an appeal to the moral basis of public life, civil disobedience is a political and not primarily a religious act. It addresses itself to the common principles of justice which men can require one another to follow and not to the aspirations of love which they cannot. Moreover by taking part in civilly disobedient acts one does not foreswear indefinitely the idea of forceful resistance; for if the appeal against injustice is repeatedly denied, then the majority has declared its intention to invite submission or resistance and the latter may conceivably be justified even in a democratic regime. We are not required to acquiesce in the crushing of fundamental liberties by democratic majorities which have shown themselves blind to the principles of justice upon which justification of the constitution depends.

THE JUSTIFICATION OF CIVIL DISOBEDIENCE

So far we have said nothing about the justification of civil disobedience, that is, the conditions under which civil disobedience may be engaged in consistent with the principles of justice that support a democratic regime. Our task is to see how the characterization of civil disobedience as addressed to the sense of justice of the majority (or to the citizens as a body) determines when such action is justified.

First of all, we may suppose that the normal political appeals to the majority have already been made in good faith and have been rejected, and that the standard means of redress have been tried. Thus, for example, existing political parties are indifferent to the claims of the minority and attempts to repeal the laws protested have been met with further repression since legal institutions are in the control of the majority. While civil disobedience should be recognized, I think, as a form of political action within the limits of fidelity to the rule of law, at the same time it is a rather desperate act just within these limits, and therefore it should, in general, be undertaken as a last resort when standard democratic processes have failed. In this sense it is not a normal political action. When it is justified there has been a serious breakdown; not only is there grave injustice in the law but a refusal more or less deliberate to correct it.

Second, since civil disobedience is a political act addressed to the sense of justice of the majority, it should usually be limited to substantial and clear violations of justice and preferably to those which, if rectified, will establish a basis for doing away with remaining injustices. For this reason there is a presumption in favor of restricting civil disobedience to violations of the first principle of justice, the principle of equal liberty, and to barriers which contravene the second principle, the principle of open offices which protects equality of opportunity. It is not, of course, always easy to tell whether these principles are satisfied. . . . When minorities are denied the right to vote or to hold certain political offices, when certain religious groups are repressed and others denied equality of opportunity in the economy, this is often obvious and there is no doubt that justice is not being given. . . .

Third, civil disobedience should be restricted to those cases where the dissenter is willing to affirm that everyone else similarly subjected to the same degree of injustice has the right to protest in a similar way. That is, we must be prepared to authorize others to dissent in similar situations and in the same way, and to accept the consequences of their doing so. Thus, we may hold, for example, that the widespread disposition to disobey civilly clear violations of fundamental liberties more or less deliberate over an extended period of time would raise the degree of justice throughout society and would insure men's self-esteem as well as their respect for one another. Indeed, I believe this to be true, though certainly it is partly a matter of conjecture. As the contract doctrine emphasizes, since the principles of justice are principles which we would agree to in an original position of equality when we do not know our social position and the like, the refusal to grant justice is either the denial of the other as an equal (as one in regard to whom we are prepared to constrain our actions by principles which we would consent to) or the manifestation of a willingness to take advantage of natural contingencies and social fortune at his expense. In either case, injustice invites submission or resistance; but submission arouses the contempt of the oppressor and confirms him in his intention. If straightway, after a decent

period of time to make reasonable political appeals in the normal way, men were in general to dissent by civil disobedience from infractions of the fundamental equal liberties, these liberties would, I believe, be more rather than less secure. Legitimate civil disobedience properly exercised is a stabilizing device in a constitutional regime, tending to make it more firmly just. . . .

The final condition, of a different nature, is the following. We have been considering when one has a right to engage in civil disobedience, and our conclusion is that one has this right should three conditions hold: when one is subject to injustice more or less deliberate over an extended period of time in the face of normal political protests; where the injustice is a clear violation of the liberties of equal citizenship; and provided that the general disposition to protest similarly in similar cases would have acceptable consequences. These conditions are not, I think, exhaustive but they seem to cover the more obvious points; yet even when they are satisfied and one has the right to engage in civil disobedience, there is still the different question of whether one should exercise this right, that is, whether by doing so one is likely to further one's ends. Having established one's right to protest one is then free to consider these tactical questions. We may be acting within our rights but still foolishly if our action only serves to provoke the harsh retaliation of the majority; and it is likely to do so if the majority lacks a sense of justice, or if the action is poorly timed or not well designed to make the appeal to the sense of justice effective. . . .

CONCLUSION: SEVERAL OBJECTIONS CONSIDERED

. . . It is natural to object to this view of civil disobedience that it relies too heavily upon the existence of a sense of justice. Some may hold that the feeling for justice is not a vital political force, and that what moves men are various other interests, the desire for wealth, power, prestige, and so on. Now this is a large question the answer to which is highly conjectural and each tends to have his own opinion. But there are two remarks which may clarify what I have said: first, I have assumed that there is in a constitutional regime a common sense of justice the principles of which are recognized to support the constitution and to guide its interpretation. In any given situation particular men may be tempted to violate these principles, but the collective force in their behalf is usually effective since they are seen as the necessary terms of cooperation among free men; and presumably the citizens of a democracy (or sufficiently many of them) want to see justice done. Where these assumptions fail, the justifying conditions for civil disobedience (the first three) are not affected, but the rationality of engaging in it certainly is. In this case, unless the costs of repressing civil dissent injures the economic self-interest (or whatever) of the majority, protest may simply make the position of the minority worse. No doubt as a tactical matter civil disobedience is more effective when its appeal coincides with other interests, but a constitutional regime is not viable in the long run without an attachment to the principles of justice of the sort which we have assumed.

Then, further, there may be a misapprehension about the manner in which a sense of justice manifests itself. There is a tendency to think that it is shown by professions of the relevant principles together with actions of an altruistic nature requiring a considerable degree of self-sacrifice. But these conditions are obviously too strong, for the majority's sense of justice may show itself simply in its being unable to undertake the measures required to suppress the minority and to punish as the law requires the various acts of civil disobedience. The sense of justice undermines the will to uphold unjust institutions, and so a majority despite its superior power may give way. It is unprepared to force the minority to be subject to injustice. Thus, although the majority's action is reluctant and grudging, the role of the sense of justice is nevertheless essential, for without it the majority would have been willing to enforce the law and to defend its position. Once we see the sense of justice as working in this negative way to make established injustices indefensible, then it is recognized as a central element of democratic politics.

Finally, it may be objected against this account that it does not settle the question of who is to say when the situation is such as to justify civil disobedience. And because it does not answer this question, it invites anarchy by encouraging

every man to decide the matter for himself. Now the reply to this is that each man must indeed settle this question for himself, although he may, of course, decide wrongly. This is true on any theory of political duty and obligation, at least on any theory compatible with the principles of a democratic constitution. The citizen is responsible for what he does. If we usually think that we should comply with the law, this is because our political principles normally lead to this conclusion. There is a presumption in favor of compliance in the absence of good reasons to the contrary. But because each man is responsible and must decide for himself as best he can whether the circumstances justify civil disobedience, it does not follow that he may decide as he pleases. It is not by looking to our personal interests or to political allegiances narrowly construed, that we should make up our mind. The citizen must decide on the basis of the principles of justice that underlie and guide the interpretation of the constitution and in the light of his sincere conviction as to how these principles should be applied in the circumstances. If he concludes that conditions obtain which justify civil disobedience and conducts himself accordingly, he has acted conscientiously and perhaps mistakenly, but not in any case at his convenience.

In a democratic society each man must act as he thinks the principles of political right require him to. We are to follow our understanding of these principles, and we cannot do otherwise. There can be no morally binding legal interpretation of these principles, not even by a supreme court or legislature. Nor is there any infallible procedure for determining what or who is right. In our system the Supreme Court, Congress, and the President often put forward rival interpretations of the Constitution. Although the Court has the final say in settling any particular case, it is not immune from powerful political influence that may change its reading of the law of the land. The Court presents its point of view by reason and argument; its conception of the Constitution must, if it is to endure, persuade men of its soundness. The final court of appeal is not the Court, or Congress, or the President, but the electorate as a whole. The civilly disobedient appeal in effect to this body. There is no danger of anarchy as long as there is a sufficient working agreement in men's conceptions of political justice and what it requires. That men can achieve such an understanding when the essential political liberties are maintained is the assumption implicit in democratic institutions. There is no way to avoid entirely the risk of divisive strife. But if legitimate civil disobedience seems to threaten civil peace, the responsibility falls not so much on those who protest as upon those whose abuse of authority and power justifies such opposition.

REVIEW AND DISCUSSION QUESTIONS

1. Why do you think Rawls calls his social-contract theory "justice as fairness"?
2. Why does Rawls think that the basic principles used to select a constitution and laws governing citizens' rights and opportunities should be chosen behind a "veil of ignorance"?
3. "Even a just constitution can lead to unjust laws." Explain using the analogy of a criminal trial.
4. Rawls thinks that the obligation to obey the law (assuming a just constitution) rests on the duty of fair play. Explain what he means.
5. How does Rawls define civil disobedience? What role does civil disobedience play in a constitutional democracy?
6. Under what circumstances does Rawls think civil disobedience is justified?
7. How does Rawls respond to those who suggest his view is naive because it is based on the view that civil disobedience is an appeal to a majority's "sense of justice"?

On Not Prosecuting Civil Disobedience

Ronald Dworkin

Besides legal obligation, civil disobedience raises further philosophical questions about the nature of law itself and whether civil disobedients should be prosecuted. Ronald Dworkin begins with a familiar argument, that since society could not function if everybody were allowed to pick and choose the laws that he or she wishes to obey, *fairness* demands uniform prosecution of all acts of civil disobedience. But in order to assess the argument, Dworkin maintains, we must first ask ourselves what we expect fellow citizens to do when they sincerely believe a law is unjust or immoral. In the American constitutional system, according to Dworkin, law and morality are not easily separated, which means that the validity of law is itself often an issue in civil disobedience cases and therefore that officials should think of civil disobedients in a different light than ordinary criminals. Ronald Dworkin teaches philosophy and law at London University and New York University.

How should the government deal with those who disobey the draft laws out of conscience? Many people think the answer is obvious: the government must prosecute the dissenters, and if they are convicted it must punish them. Some people reach this conclusion easily, because they hold the mindless view that conscientious disobedience is the same as lawlessness. They think that the dissenters are anarchists who must be punished before their corruption spreads. Many lawyers and intellectuals come to the same conclusion, however, on what looks like a more sophisticated argument. They recognize that disobedience to law may be *morally* justified, but they insist that it cannot be *legally* justified, and they think that it follows from this truism that the law must be enforced. Erwin Griswold, the Solicitor General of the United States, and the former dean of the Harvard Law School, appears to have adopted this view in a recent statement. "[It] is of the essence of law," he said, "that it is equally applied to all, that it binds all alike, irrespective of personal motive. For this reason, one who contemplates civil disobedience out of moral conviction should not be surprised and must not be bitter if a criminal conviction ensues. And he must accept the fact that organized society cannot endure on any other basis." . . .

But the argument that, because the government believes a man has committed a crime, it must prosecute him is much weaker than it seems. Society "cannot endure" if it tolerates all disobedience; it does not follow, however, nor is there evidence, that it will collapse if it tolerates some. In the United States prosecutors have discretion whether to enforce criminal laws in particular cases. A prosecutor may properly decide not to press charges if the lawbreaker is young, or inexperienced, or the sole support of the family, or is repentant, or turns state's evidence, or if the law is unpopular or unworkable or generally disobeyed, or if the courts are clogged with more important cases, or for dozens of other reasons. This discretion is not license—we expect prosecutors to have good reasons for exercising it—but there are, at least *prima facie*, some good reasons for not prosecuting those who disobey the draft laws out of conscience. One is the obvious reason that they act out of better motives than those who break the law out of greed or a desire to subvert government. Another is the practical reason that our society suffers a loss if it punishes a group that includes—as a group of draft dissenters does—some of its most thoughtful and loyal citizens. Jailing such men solidifies their alienation from society, and alienates many like them who are deterred by the threat.

Those who think that conscientious draft offenders . . . should always be punished must show that these are not good reasons for exercising discretion, or they must find contrary reasons that outweigh them. What arguments might they produce? . . . Dean Griswold and those who agree with him seem to rely on a fundamental moral argument that it would be unfair, not merely impractical, to let the dissenters go unpunished. They think it would be unfair, I gather, because

society could not function if everyone disobeyed laws he disapproved of or found disadvantageous. If the government tolerates those few who will not "play the game," it allows them to secure the benefits of everyone else's deference to law, without shouldering the burdens, such as the burden of the draft.

This argument is a serious one. It cannot be answered simply by saying that the dissenters would allow everyone else the privilege of disobeying a law he believed immoral. In fact, few draft dissenters would accept a changed society in which sincere segregationists were free to break civil rights laws they hated. The majority want no such change, in any event, because they think that society would be worse off for it; until they are shown this is wrong, they will expect their officials to punish anyone who assumes a privilege which they, for the general benefit, do not assume.

There is, however, a flaw in the argument. The reasoning contains a hidden assumption that makes it almost entirely irrelevant to the draft cases, and indeed to any serious case of civil disobedience in the United States. The argument assumes that the dissenters know that they are breaking a valid law, and that the privilege they assert is the privilege to do that. Of course, almost everyone who discusses civil disobedience recognizes that in America a law may be invalid because it is unconstitutional. . . .

Doubtful law is by no means special or exotic in cases of civil disobedience. On the contrary. In the United States, at least, almost any law which a significant number of people would be tempted to disobey on moral grounds would be doubtful—if not clearly invalid—on constitutional grounds as well. The constitution makes our conventional political morality relevant to the question of validity; any statute that appears to compromise that morality raises constitutional questions, and if the compromise is serious, the constitutional doubts are serious also.

[*Editors' Note: Dworkin illustrates this by pointing out that moral objections to the Vietnam War—for example, that the United States was using immoral weapons, that Congress had never officially declared war, and that it had been made a crime to "counsel" draft resistance—have parallels in law, for example, international treaties we signed banning certain weapons, Congress's exclusive power to declare war, and First Amendment protections of free speech. These moral/legal arguments thus provide the basis for thinking that, in fact, the laws in question are invalid.*]

. . . We cannot conclude from these arguments that the draft (or any part of it) was unconstitutional. . . . But the arguments of unconstitutionality were at least plausible and a reasonable and competent lawyer might think that they present a stronger case, on balance, than the counter arguments. . . .

Therefore we cannot assume, in judging what should have been done with the draft dissenters, that they were asserting a privilege to disobey valid laws. We cannot decide that fairness demanded their punishment until we try to answer further questions: What should a citizen do when the law is unclear, and when he thinks it allows what others think it does not? I do not mean to ask, of course, what it is *legally* proper for him to do, or what his *legal* rights are—that would be begging the question, because it depends upon whether he is right or they are right. I mean to ask what his proper course is as a citizen, what, in other words, we would consider to be "playing the game." That is a crucial question, because it cannot be unfair not to punish him if he is acting as, given his opinions, we think he should.[1]

There is no obvious answer on which most citizens would readily agree, and that is itself significant. If we examine our legal institutions and practices, however, we shall discover some relevant underlying principles and policies. I shall set out three possible answers to the question, and then try to show which of these best fits our practices and expectations. The three possibilities I want to consider are these:

1. If the law is doubtful, and it is therefore unclear whether it permits someone to do what he wants, he should assume the worst, and act on the assumption that it does not. He should obey the executive authorities who command him, even though he thinks they are wrong, while using the political process, if he can, to change the law.

2. If the law is doubtful, he may follow his own judgment, that is, he may do what he wants if he believes that the case that the law permits this is stronger than the case that it does not. But he may

follow his own judgment only until an authoritative institution, like a court, decides the other way in a case involving him or someone else. Once an institutional decision has been reached, he must abide by that decision, even though he thinks that it was wrong. (There are, in theory, many subdivisions of this second possibility. We may say that the individual's choice is foreclosed by the contrary decision of any court, including the lowest court in the system if the case is not appealed. Or we may require a decision of some particular court or institution. I shall discuss this second possibility in its most liberal form, namely that the individual may properly follow his own judgment until a contrary decision of the highest court competent to pass on the issue, which, in the case of the draft, was the United States Supreme Court.)

3. If the law is doubtful, he may follow his own judgment, even after a contrary decision by the highest competent court. Of course, he must take the contrary decision of any court into account in making his judgment of what the law requires. Otherwise the judgment would not be an honest or reasonable one, because the doctrine of precedent, which is an established part of our legal system, has the effect of allowing the decision of the courts to change the law. Suppose, for example, that a taxpayer believes that he is not required to pay tax on certain forms of income. If the Supreme Court decides to the contrary, he should, taking into account the practice of according great weight to the decisions of the Supreme Court on tax matters, decide that the Court's decision has itself tipped the balance, and that the law now requires him to pay the tax.

Someone might think that this qualification erases the difference between the third and the second models, but it does not. The doctrine of precedent gives different weights to the decisions of different courts, and greatest weight to the decisions of the Supreme Court, but it does not make the decisions of any court conclusive. Sometimes, even after a contrary Supreme Court decision, an individual may still reasonably believe that the law is on his side; such cases are rare, but they are most likely to occur in disputes over constitutional law when civil disobedience is involved. The Court has shown itself more likely to overrule its past decisions if these have limited important personal or political rights, and it is just these decisions that a dissenter might want to challenge.

We cannot assume, in other words, that the Constitution is always what the Supreme Court says it is. Oliver Wendell Holmes, for example, did not follow such a rule in his famous dissent in the *Gitlow* case. A few years before, in *Abrams*, he had lost his battle to persuade the court that the First Amendment protected an anarchist who had been urging general strikes against the government. A similar issue was presented in *Gitlow*, and Holmes once again dissented. "It is true," he said, "that in my opinion this criterion was departed from [in *Abrams*] but the convictions that I expressed in that case are too deep for it to be possible for me as yet to believe that it . . . settled the law." Holmes voted for acquitting Gitlow, on the ground that what Gitlow had done was no crime, even though the Supreme Court had recently held that it was.

Here then are three possible models for the behavior of dissenters who disagree with the executive authorities when the law is doubtful. Which of them best fits our legal and social practices?

I think it plain that we do not follow the first of these models, that is, that we do not expect citizens to assume the worst. If no court has decided the issue, and a man thinks, on balance, that the law is on his side, most of our lawyers and critics think it perfectly proper for him to follow his own judgment. Even when many disapprove of what he does—such as peddling pornography—they do not think he must desist just because the legality of his conduct is subject to doubt.

It is worth pausing a moment to consider what society would lose if it did follow the first model or, to put the matter the other way, what society gains when people follow their own judgment in cases like this. When the law is uncertain, in the sense that lawyers can reasonably disagree on what a court ought to decide, the reason usually is that different legal principles and policies have collided, and it is unclear how best to accommodate these conflicting principles and policies.

Our practice, in which different parties are encouraged to pursue their own understanding, provides a means of testing relevant hypotheses. If the question is whether a particular rule would have certain undesirable consequences, or whether these consequences would have limited or broad

ramifications, then, before the issue is decided, it is useful to know what does in fact take place when some people proceed on that rule. (Much anti-trust and business regulation law has developed through this kind of testing.) If the question is whether and to what degree a particular solution would offend principles of justice or fair play deeply respected by the community, it is useful, again, to experiment by testing the community's response. The extent of community indifference to anti-contraception laws, for example, would never have become established had not some organizations deliberately flouted those laws.

If the first model were followed, we would lose the advantages of these tests. The law would suffer, particularly if this model were applied to constitutional issues. When the validity of a criminal statute is in doubt, the statute will almost always strike some people as being unfair or unjust, because it will infringe some principle of liberty or justice or fairness which they take to be built into the Constitution. If our practice were that whenever a law is doubtful on these grounds, one must act as if it were valid, then the chief vehicle we have for challenging the law on moral grounds would be lost, and over time the law we obeyed would certainly become less fair and just, and the liberty of our citizens would certainly be diminished. . . .

We must also reject the second model, that if the law is unclear a citizen may properly follow his own judgment until the highest court has ruled that he is wrong. This fails to take into account the fact that any court, including the Supreme Court, may overrule itself. In 1940 the Court decided that a West Virginia law requiring students to salute the Flag was constitutional. In 1943 it reversed itself, and decided that such a statute was unconstitutional after all. What was the duty as citizens, of those people who in 1941 and 1942 objected to saluting the Flag on grounds of conscience, and thought that the Court's 1940 decision was wrong? We can hardly say that their duty was to follow the first decision. They believed that saluting the Flag was unconscionable, and they believed, reasonably, that no valid law required them to do so. The Supreme Court later decided that in this they were right. The Court did not simply hold that after the second decision failing to salute would not be a crime; it held (as in a case like this it almost always would) that it was no crime after the first decision either.

Some will say that the flag-salute dissenters should have obeyed the Court's first decision, while they worked in the legislatures to have the law repealed, and tried in the courts to find some way to challenge the law again without actually violating it. That would be, perhaps, a plausible recommendation if conscience were not involved, because it would then be arguable that the gain in orderly procedure was worth the personal sacrifice of patience. But conscience was involved, and if the dissenters had obeyed the law while biding their time, they would have suffered the irreparable injury of having done what their conscience forbade them to do. It is one thing to say that an individual must sometimes violate his conscience when he knows that the law commands him to do it. It is quite another to say that he must violate his conscience even when he reasonably believes that the law does not require it, because it would inconvenience his fellow citizens if he took the most direct, and perhaps the only, method of attempting to show that he is right and they are wrong.

Since a court may overrule itself, the same reasons we listed for rejecting the first model count against the second as well. If we did not have the pressure of dissent, we would not have a dramatic statement of the degree to which a court decision against the dissenter is felt to be wrong, a demonstration that is surely pertinent to the question of whether it was right. We would increase the chance of being governed by rules that offend the principles we claim to serve. . . .

Thus the third model, or something close to it, seems to be the fairest statement of a man's social duty in our community. A citizen's allegiance is to the law, not to any particular person's view of what the law is, and he does not behave unfairly so long as he proceeds on his own considered and reasonable view of what the law requires. Let me repeat (because it is crucial) that this is not the same as saying that an individual may disregard what the courts have said. The doctrine of precedent lies near the core of our legal system, and no one can make a reasonable effort to follow the law unless he grants the courts the general power to alter it by their decisions. But if the issue is one touching fundamental personal or political rights, and it is

arguable that the Supreme Court has made a mistake, a man is within his social rights in refusing to accept that decision as conclusive. . . .

. . . I have been talking about the case of a man who believes that the law is not what other people think, or what the courts have held. This description may fit some of those who disobey the draft laws out of conscience, but it does not fit most of them. Most of the dissenters are not lawyers or political philosophers; they believe that the laws on the books are immoral, and inconsistent with their country's legal ideals, but they have not considered the question of whether they may be invalid as well. Of what relevance to their situation, then, is the proposition that one may properly follow one's own view of the law?

To answer this, I shall have to return to the point I made earlier. The Constitution, through the due process clause [government cannot deprive citizens of "life, liberty, or property without due process of law"], the equal protection clause [states cannot deny citizens the "equal protection of the laws"], the First Amendment [government cannot "abridge freedom of speech" or make a law "respecting the establishment of religion, or prohibiting the free exercise thereof"], and the other provisions I mentioned, injects an extraordinary amount of our political morality into the issue of whether a law is valid. The statement that most draft dissenters are unaware that the law is invalid therefore needs qualification. They hold beliefs that, if true, strongly support the view that the law is on their side; the fact that they have not reached that further conclusion can be traced, in at least most cases, to their lack of legal sophistication. If we believe that when the law is doubtful people who follow their own judgment of the law may be acting properly, it would seem wrong not to extend that view to those dissenters whose judgments come to the same thing. No part of the case that I made for the third model would entitle us to distinguish them from their more knowledgeable colleagues.

We can draw several tentative conclusions from the argument so far: When the law is uncertain, in the sense that a plausible case can be made on both sides, then a citizen who follows his own judgment is not behaving unfairly. Our practices permit and encourage him to follow his own judgment in such cases. For that reason, our government has a special responsibility to try to protect him, and soften his predicament, whenever it can do so without great damage to other policies. It does not follow that the government can guarantee him immunity—it cannot adopt the rule that it will prosecute no one who acts out of conscience, or convict no one who reasonably disagrees with the courts. That would paralyze the government's ability to carry out its policies; it would, moreover, throw away the most important benefit of following the third model. If the state never prosecuted, then the courts could not act on the experience and the arguments the dissent has generated. But it does follow that when the practical reasons for prosecuting are relatively weak in a particular case, or can be met in other ways, the path of fairness lies in tolerance. The popular view that the law is the law and must always be enforced refuses to distinguish the man who acts on his own judgment of a doubtful law, and thus behaves as our practices provide, from the common criminal. I know of no reason, short of moral blindness, for not drawing a distinction in principle between the two cases.

I anticipate a philosophical objection to these conclusions: that I am treating law as a "brooding omnipresence in the sky." I have spoken of people making judgments about what the law requires, even in cases in which the law is unclear and undemonstrable. I have spoken of cases in which a man might think that the law requires one thing, even though the Supreme Court has said that it requires another, and even when it was not likely that the Supreme Court would soon change its mind. I will therefore be charged with the view that there is always a "right answer" to a legal problem to be found in natural law or locked up in some transcendental strongbox.

The strongbox theory of law is, of course, nonsense. When I say that people hold views on the law when the law is doubtful, and that these views are not merely predictions of what the courts will hold, I intend no such metaphysics. I mean only to summarize as accurately as I can many of the practices that are part of our legal process.

Lawyers and judges make statements of legal right and duty, even when they know these are not demonstrable, and support them with arguments even when they know that these arguments will not

appeal to everyone. They make these arguments to one another, in the professional journals, in the classrooms, and in the courts. They respond to these arguments, when others make them, by judging them good or bad or mediocre. In so doing they assume that some arguments for a given doubtful position are better than others. They also assume that the case on one side of a doubtful proposition may be stronger than the case on the other, which is what I take a claim of law in a doubtful case to mean. . . .

Our legal system pursues these goals (the development and testing of the law through experimentation by citizens and through the adversary process) by inviting citizens to decide the strengths and weaknesses of legal arguments for themselves, or through their own counsel, and to act on these judgments, although that permission is qualified by the limited threat that they may suffer if the courts do not agree. Success in this strategy depends on whether there is sufficient agreement within the community on what counts as a good or bad argument, so that, although different people will reach different judgments, these differences will be neither so profound nor so frequent as to make the system unworkable, or dangerous for those who act by their own lights. I believe there is sufficient agreement on the criteria of the argument to avoid these traps, although one of the main tasks of legal philosophy is to exhibit and clarify these criteria. In any event, the practices I have described have not yet been shown to be misguided; they therefore must count in determining whether it is just and fair to be lenient to those who break what others think is the law.

I have said that the government has a special responsibility to those who act on a reasonable judgment that a law is invalid. It should make accommodation for them as far as possible, when this is consistent with other policies. It may be difficult to decide what the government ought to do, in the name of that responsibility, in particular cases. The decision will be a matter of balance, and flat rules will not help. . . .

Some lawyers will be shocked by my general conclusion that we have a responsibility toward those who disobey the draft laws out of conscience, and that we may be required not to prosecute them, but rather to change our laws or adjust our sentencing procedures to accommodate them. The simple Draconian propositions, that crime must be punished, and that he who misjudges the law must take the consequences, have an extraordinary hold on the professional as well as the popular imagination. But the rule of law is more complex and more intelligent than that and it is important that it survive.

NOTE

1. I do not mean to imply that the government should always punish a man who deliberately breaks a law he knows is valid. There may be reasons of fairness or practicality, like those I listed in the second paragraph, for not prosecuting such men. But cases like the draft cases present special arguments for tolerance; I want to concentrate on these arguments and therefore have isolated these cases.

REVIEW AND DISCUSSION QUESTIONS

1. Describe how Dworkin understands the relationship between morality and law. Why is it often difficult to know whether a law is valid?
2. What service does Dworkin think civil disobedients provide in our system?
3. Dworkin argues that it is not unfair to refuse to prosecute civil disobedients. His argument relies on a discussion of three courses of action open to citizens who object to law, and what we as a society reasonably expect of our fellow citizens. Explain his argument.
4. What does Dworkin mean in rejecting the "strongbox" view of law? If law and morality are connected, as he suggests, how is the law to be tested and legal arguments assessed?
5. In light of one or two of the cases you have read, is Dworkin's description of the connections between law and morality an accurate one?

PART TWO

Legal Reasoning

CHAPTER 4

Statutes

Legal reasoning has various aspects, but much of it involves the interpretation of statutes and precedents. Statutes are written in explicit language and adopted by a legislative body, whether the U.S. Congress, a state legislature, or a local county or city council. A statute can be as simple as "no vehicles in the park" or as complex as a tax code, but none of them is self-interpreting. Readings in this section discuss some of the issues surrounding how courts interpret statutes and the justifications for the methods that are used. We begin, however, with two imaginary cases.

On Interpretation: The Adultery Clause of the Ten Commandments

Sanford Levinson

Using an ingenious example, Sanford Levinson describes a variety of problems confronting anybody who undertakes to interpret broad concepts with ancient roots, such as the Ten Commandments. Professor Levinson teaches at the University of Texas Law School.

Consider the following problem:

In 1990, a number of concerned citizens, worried about what they regarded as the corruption of American life, met to consider what could be done. During the course of the discussion, one of the speakers electrified the audience with the following comments:

> The cure for our ills is a return to old-time religion, and the best single guide remains the Ten Commandments. Whenever I am perplexed as to what I ought to do, I turn to the Commandments for the answer, and I am never disappointed. Sometimes I don't immediately like what I discover, but then I think more about the problem and realize how limited my perspective is compared to that of the framer of those great words. Indeed, all that is necessary is for everyone to obey the Ten Commandments, and our problems will all be solved.[1]

Within several hours the following plan was devised: As part of the effort to encourage a return to the "old-time religion" of the Ten Commandments, a number of young people would be asked to take an oath on their eighteenth birthday to "obey, protect, support, and defend the Ten Commandments" in all of their actions. If the person complied with the oath for seventeen years, he or she would receive an award of $10,000 on his or her thirty-fifth birthday.

The Foundation for the Ten Commandments was funded by the members of the 1990 convention, plus the proceeds of a national campaign for contributions. The speaker quoted above contributed $20 million, and an additional $30 million was collected, $15 million from the convention and $15 million from the national campaign. The interest generated by the $50 million is approximately $6 million per year. Each year since 1990, 500 persons have taken the oath. *You* are appointed sole trustee of the Foundation, and your most important duty is to determine whether the oath-takers have complied with their vows and are thus entitled to the $10,000.

It is now 2007, and the first set of claimants comes before you:

> **1.** Claimant *A* is a married male. Although freely admitting that he has had sexual intercourse with a number of women other than his wife during their marriage, he brings to your attention the fact that "adultery," at the time of Biblical Israel, referred only to the voluntary intercourse of a married woman with a man other than her husband. He specifically notes the following passage from the article *Adultery,* I Jewish Encyclopedia 314:
>
> > The extramarital intercourse of a married man is not *per se* a crime in biblical or later Jewish law. This distinction stems from the economic aspect of Israelite marriage: The wife as the husband's possession . . . , and adultery constituted a violation of the husband's exclusive right to her; the wife, as the husband's possession, had no such right to him.

A has taken great care to make sure that all his sexual partners were unmarried, and thus he claims to have been faithful to the original understanding of the Ten Commandments. However we might define "adultery" today, he argues, is irrelevant. His oath was to comply with the Ten Commandments; he claims to have done so. (It is stipulated that *A,* like all the other claimants, has complied with all the other commandments; the only question involves compliance with the commandment against adultery.)

Upon further questioning, you discover that no line-by-line explication of the Ten Commandments was proffered in 1990 at the time that *A* took the oath. But, says *A,* whenever a question arose in his mind as to what the Ten Commandments required of him, he made conscientious attempts to research the particular issue. He initially shared your (presumed) surprise at the results of his research, but further study indicated that all authorities agreed with the scholars who wrote the *Jewish Encyclopedia* regarding the original understanding of the Commandment.

> **2.** Claimant *B* is *A*'s wife, who admits that she has had extramarital relationships with other men. She notes, though, that these affairs were entered into with the consent of her husband. In response to the fact that she undoubtedly violated the ancient understanding of "adultery," she states that that understanding is fatally outdated:
>
> **a.** It is unfair to distinguish between the sexual rights of males and females. That the Israelites were outrageously sexist is no warrant for your maintaining the discrimination.
>
> **b.** Moreover, the reason for the differentiation, as already noted, was the perception of the wife as property. That notion is a repugnant one that has been properly repudiated by all rational thinkers, including all major branches of the Judeo-Christian religious tradition historically linked to the Ten Commandments.
>
> **c.** She further argues that, insofar as the modern prohibition of adultery is defensible, it rests on the ideal of discouraging deceit and the betrayal of promises of sexual fidelity. But these admittedly negative factors are not present in her case because she had scrupulously informed her husband and received his consent, as required by their marriage contract outlining the terms of their "open marriage."
>
> (It turns out, incidentally, that *A* had failed to inform his wife of at least one of his sexual encounters. Though he freely admits that this constitutes a breach of the contract he had made with *B,* he nevertheless returns to his basic argument about original understanding, which makes consent irrelevant.)

3. *C,* a male (is this relevant?), is the participant in a bigamous marriage. *C* had no sexual encounters beyond his two wives. (He also points out that bigamy was clearly tolerated in both pre- and post-Sinai Israel and indeed was accepted within the Yemenite community of Jews well into the twentieth century. It is also accepted in a variety of world cultures.)

4. *D,* a practicing Christian, admits that he has often lusted after women other than his wife. (Indeed, he confesses as well that it was only after much contemplation that he decided not to sexually consummate a relationship with a coworker whom he thinks he "may love" and with whom he has held hands.) You are familiar with Christ's words, Matthew 5:28: "Whosoever looketh on a woman to lust after, he hath committed adultery with her already in his heart." (Would it matter to you if *D* were the wife, who had lusted after other men?)

5. Finally, claimant *E* has never even lusted after another woman since his marriage on the same day he took his oath. He does admit, however, to occasional lustful fantasies about his wife. *E,* a Catholic, is shocked when informed of Pope John Paul II's statement that "adultery in your heart is committed not only when you look with concupiscence at a woman who is not your wife, but also if you look in the same manner at your wife." The Pope's rationale apparently is that all lust, even that directed toward a spouse, dehumanizes and reduces the other person "to an erotic object."

Which, if any, of the claimants should get the $10,000? . . .

NOTE

1. *Cf.* Statement of President Ronald Reagan, Press Conference, Feb. 21, 1985, reprinted in *N.Y. Times,* Feb. 22, 1985 § 1, at 10, col. 3: "I've found that the Bible contains an answer to just about everything and every problem that confronts us, and I wonder sometimes why we won't recognize that one Book could solve a lot of problems for us."

REVIEW AND DISCUSSION QUESTIONS

1. Describe the situation Levinson envisions.
2. Summarize the arguments of each of the five claimants.
3. Which of the five, if any, should get the money? Explain.
4. Would an unmarried, stable gay or lesbian couple be guilty of adultery?
5. What principles or ideals should guide judges who interpret laws? Why are those principles important for them to follow?

Interpreting the Small Bird Act

Regina v. Ojibway

In this fictional case, the issue before the Court is whether a horse is legally a bird.

Blue, J.: This is an appeal by the Crown by way of a stated case from a decision of the magistrate acquitting the accused of a charge under the Small Birds Act, R.S.O., 1960, c.724, s.2. The facts are not in dispute. Fred Ojibway, an Indian, was riding his pony through Queen's Park on January 2, 1965. Being impoverished, and having been forced to pledge his saddle, he substituted a downy pillow in lieu of the said saddle. On this particular day the accused's misfortune was further heightened by the

circumstance of his pony breaking its right foreleg. In accord with current Indian custom, the accused then shot the pony to relieve it of its awkwardness.

The accused was then charged with having breached the Small Birds Act, s.2 of which states:

> 2. Anyone maiming, injuring or killing small birds is guilty of an offense and subject to a fine not in excess of two hundred dollars.

The learned magistrate acquitted the accused, holding, in fact, that he had killed his horse and not a small bird. With respect, I cannot agree.

In light of the definition section my course is quite clear. Section 1 defines "bird" as "a two legged animal covered with feathers." There can be no doubt that this case is covered by this section.

Counsel for the accused made several ingenious arguments to which, in fairness, I must address myself. He submitted that the evidence of the expert clearly concluded that the animal in question was a pony and not a bird, but this is not the issue. We are not interested in whether the animal in question is a bird or not in fact, but whether it is one in law. Statutory interpretation has forced many a horse to eat birdseed for the rest of its life.

Counsel also contended that the neighing noise emitted by the animal could not possibly be produced by a bird. With respect, the sounds emitted by an animal are irrelevant to its nature, for a bird is no less a bird because it is silent.

Counsel for the accused also argued that since there was evidence to show accused had ridden the animal, this pointed to the fact that it could not be a bird but was actually a pony. Obviously, this avoids the issue. The issue is not whether the animal was ridden or not, but whether it was shot or not, for to ride a pony or a bird is of no offense at all. I believe that counsel now sees his mistake.

Counsel contends that the iron shoes found on the animal decisively disqualify it from being a bird. I must inform counsel, however, that how an animal dresses is of no concern to this court.

Counsel relied on the decision *In Re Chickadee,* where he contends that in similar circumstances the accused was acquitted. However, this is a horse of a different colour. A close reading of that case indicates that the animal in question there was not a small bird, but, in fact, a midget of a much larger species. Therefore, that case is inapplicable to our facts.

Counsel finally submits that the word "small" in the title Small Birds Act refers not to "Birds" but to "Act," making it The Small Act relating to Birds. With respect, counsel did not do his homework very well, for the Large Birds Act, R.S.O., 1960, c.725, is just as small. If pressed, I need only refer to the Small Loans Act, R.S.O., 1960, c.727, which is twice as large as the Large Birds Act.

It remains then to state my reason for judgment which, simply, is as follows: Different things may take on the same meaning for different purposes. For the purpose of The Small Birds Act, all two legged, feather-covered animals are birds. This, of course, does not imply that only two legged animals qualify, for the legislative intent is to make two legs merely the minimum requirement. The statute therefore contemplated multi-legged animals with feathers as well. Counsel submits that having regard to the purpose of the statute only small animals "naturally covered" with feathers could have been contemplated. However, had this been the intention of the legislature, I am certain that the phrase "naturally covered" would have been expressly inserted just as "Long" was inserted in the Longshoreman's Act.

Therefore, a horse with feathers on its back must be deemed for the purposes of this Act to be a bird, and *a fortiori,* a pony with feathers on its back is a small bird.

Counsel posed the following rhetorical question: If the pillow had been removed prior to the shooting, would the animal still be a bird? To this let me answer rhetorically: Is a bird any less of a bird without its feathers?

Appeal allowed.

REVIEW AND DISCUSSION QUESTIONS

1. What are the relevant facts and law in this case?
2. Discuss what is wrong with the "reasoning" of the judge in this case, in light of the principles or ideals that you believe should guide a judge charged with interpreting a law.

A Case Study in Interpretation: The Mann Act

Edward H. Levi

In this essay, Edward H. Levi begins by distinguishing case-law reasoning from statutory interpretation. Central to deciding what a statute means is the question of the intention of the legislature. He then illustrates these problems by describing the legislative history of the Mann Act (sometimes called the "White Slave Traffic Act"). Edward H. Levi was professor of law at the University of Chicago.

It is customary to think of case-law reasoning as inductive and the application of statutes as deductive. The thought seems erroneous but the emphasis has some meaning. With case law the concepts can be created out of particular instances. . . .

The application of a statute seems to be in great contrast. The words are given. They are not to be taken lightly since they express the will of the legislature. The legislature is the law-making body. It looks like deduction to apply the word to the specific case.

. . . The legislature may have had a particular case uppermost in mind, but it has spoken in general terms. Not only respect but application is due to the general words the legislature used. The rules for statutory construction make the same point. They are words which tell one how to operate a given classification system. The problem is to place the species inside the genus and the particular case inside the species. The words used by the legislature are treated as words of classification which are to be applied. Yet the rules [of construction] themselves show that there may be some ambiguity in the words used. The words are to be construed in the light of the meaning given to other words in the same or related statute. The specification of particular instances indicates that similar but unmentioned instances are not to be included. But the specification of particular instances, when in addition a word of a general category is used, may be the indication that other like instances are also intended.

Thus in the application of a statute the intent of the legislature seems important. The rules of construction are ways of finding out the intent. The actual words used are important but insufficient. The report of congressional committees may give some clue. Prior drafts of the statute may show where meaning was intentionally changed. Bills presented but not passed may have some bearing. Words spoken in debate may now be looked at. Even the conduct of the litigants may be important in that the failure of the government to have acted over a period of time on what it now suggests as the proper interpretation throws light on the common meaning. But it is not easy to find the intent of the legislature.

Justice Reed has given us some Polonius-sounding advice on the matter:

> There is, of course, no more persuasive evidence of the purpose of a statute than the words by which the legislature undertook to give expression to its wishes. Often these words are sufficient in and of themselves to determine the purpose of the legislature. In such cases we have followed their plain meaning. When that meaning had led to absurd or futile results, however, this Court has looked beyond the words to the purpose of the act. Frequently, however, even when the plain meaning did not produce absurd results but merely an unreasonable one "plainly at variance with the policy of legislation as a whole" this Court has followed that purpose rather than the literal words. . . . Obviously there is danger that the courts' conclusion as to legislative purpose will be unconsciously influenced by the judges' own views or by factors not considered by the enacting body. A lively appreciation of the danger is the best assurance of escape from its threat but hardly justifies an acceptance of a literal interpretation dogma which withholds from the courts available information for reaching a correct conclusion. Emphasis should be laid too upon the necessity for appraisal of the purposes as a whole of Congress in analyzing the meaning of clauses of sections of general acts. A few words of general connotation appearing in the text of statutes should not be given a wide meaning, contrary to settled policy, "except as a different purpose is plainly shown."[1]

The words of advice force one to re-examine whether there is any difference between case law and statutory interpretation. It is not enough to show that the words used by the legislature have some meaning. Concepts created by case law also have some meaning, but the meaning is ambiguous. It is not clear how wide or narrow the scope is to be. Can it be said that the words used by the legislature have any more meaning than that, or is there the same ambiguity? One important difference can be noted immediately. Where case law is considered, there is a conscious realignment of cases; the problem is not the intention of the prior judge. But with a statute the reference is to the kind of things intended by the legislature. All concepts suggest, but case-law concepts can be re-worked. A statutory concept, however, is supposed to suggest what the legislature had in mind; the items to be included under it should be of the same order. We mean to accomplish what the legislature intended. This is what Justice Reed has said. The difficulty is that what the legislature intended is ambiguous. In a significant sense there is only a general intent which preserves as much ambiguity in the concept used as though it had been created by case law.

This is not the result of inadequate draftsmanship, as is so frequently urged. Matters are not decided until they have to be. For a legislature perhaps the pressures are such that a bill has to be passed dealing with a certain subject. But the precise effect of the bill is not something upon which the members have to reach agreement. If the legislature were a court, it would not decide the precise effect until a specific fact situation arose demanding an answer. Its first pronouncement would not be expected to fill in the gaps. But since it is not a court, this is even more true. It will not be required to make the determination in any event, but can wait for the court to do so. There is a related and an additional reason for ambiguity. As to what type of situation is the legislature to make a decision? Despite much gospel to the contrary, a legislature is not a fact-finding body. There is no mechanism, as there is with a court, to require the legislature to sift facts and to make a decision about specific situations. There need be no agreement about what the situation is. The members of the legislative body will be talking about different things; they cannot force each other to accept even a hypothetical set of facts. The result is that even in a non-controversial atmosphere just exactly what has been decided will not be clear.

Controversy does not help. Agreement is then possible only through escape to a higher level of discourse with greater ambiguity. This is one element which makes compromise possible. Moreover, from the standpoint of the individual member of the legislature there is reason to be deceptive. He must escape from pressures at home. Newspapers may have created an atmosphere in which some legislation must be passed. Perhaps the only chance to get legislation through is to have it mean something not understood by some colleagues. . . . And if all this were not sufficient, it cannot be forgotten that to speak of legislative intent is to talk of group action, where much of the group may be ignorant or misinformed. Yet the emphasis should not be on this fact, but on the necessity that there be ambiguity before there can be any agreement about how unknown cases will be handled.

But the court will search for the legislative intent, and this does make a difference. Its search results in an initial filling-up of the gap. The first opinions may not definitely set the whole interpretation. A more decisive view may be edged toward, but finally there is likely to be an interpretation by the court which gives greater content to the words used. In building up this interpretation, the reference will be to the kind of examples that the words used, as commonly understood, would call to mind. Reasoning by example will then proceed from that point. There is a difference then from case law in that the legislature has compelled the use of one word. The word will not change verbally. It could change in meaning, however, and if frequent appeals as to what the legislature really intended are permitted, it may shift radically from time to time. When this is done, a court in interpreting legislation has really more discretion than it has with case law. For it can escape from prior cases by saying that they have ignored the legislative intent.

There is great danger in this. Legislatures and courts are cooperative law-making bodies. It is important to know where the responsibility lies. If legislation which is disfavored can be interpreted away from time to time, then it is not to be expected, particularly if controversy is high, that

the legislature will ever act. It will always be possible to say that new legislation is not needed because the court in the future will make a more appropriate interpretation. If the court is to have freedom to reinterpret legislation, the result will be to relieve the legislature from pressure. The legislation needs judicial consistency. Moreover, the court's own behavior in the face of pressure is likely to be indecisive. In all likelihood it will do enough to prevent legislative revision and not much more. Therefore it seems better to say that once a decisive interpretation of legislative intent has been made, and in that sense a direction has been fixed within the gap of ambiguity, the court should take that direction as given. In this sense a court's interpretation of legislation is not dictum. The words it uses do more than decide the case. They give broad direction to the statute.

The doctrine which is suggested here is a hard one. In many controversial situations, legislative revision cannot be expected. It often appears that the only hope lies with the courts. Yet the democratic process seems to require that controversial changes should be made by the legislative body. This is not only because there is a mechanism for holding legislators responsible. It is also because courts are normally timid. Since they decide only the case before them, it is difficult for them to compel any controversial reform unless they are willing to hold to an unpopular doctrine over a sustained period of time. The difficulties which administrative agencies have in the face of sustained pressure serve as a warning. When courts enter the area of great controversy, they require unusual protection. They must be ready to appeal to the constitution.

Where legislative interpretation is concerned, therefore, it appears that legal reasoning does attempt to fix the meaning of the word. When this is done, subsequent cases must be decided upon the basis that the prior meaning remains. It must not be re-worked. Its meaning is made clear as examples are seen, but the reference is fixed. It is a hard doctrine against which judges frequently rebel. The Mann Act is a good example.

On June 25, 1910, the Mann Act, which recites that it "shall be known and referred to as the 'White Slave Traffic Act,'" went into effect. The Act provides in part: "Any person who shall knowingly transport or cause to be transported, or aid or assist in obtaining transportation for, or in transporting, in interstate or foreign commerce or in any territory or in the District of Columbia, any woman or girl for the purpose of prostitution or debauchery, or for any other immoral purpose, or with the intent and purpose to induce, entice or compel such woman or girl to become a prostitute, or to give herself up to debauchery, or to engage in any other immoral practice . . . shall be deemed guilty of a felony." The Act was not passed in haste. Indeed, the matter was much debated and prior reports about it had been written. The Secretary of Commerce and Labor had discussed the problem in his 1908 report; so had an Immigration Commission in a preliminary report for 1909. There were international aspects to the problem, and a treaty had been concluded. The President had directed the attention of Congress to the need for legislation, and the proposed bill had been considered in majority and minority congressional committee reports.

The Mann Act was passed during a period when large American cities had illegal but segregated "red-light" areas. It was believed that women were procured for houses of prostitution by bands of "white slavers" who "were said to operate from coast to coast, in town and country, with tentacles in foreign lands, east and west and across the American borders. The most sensational of these were said to be the French, Italian, and Jewish rings who preyed on innocent girls of their respective nationalities at ports of entry into the United States or ensnared them at the ports of embarkation in Europe and even in their home towns." It was thought that the girls were young; many of them were supposed to be scarcely in their teens. They were forced or lured into the business. It was thought that they had previously been virtuous, and while supposedly many of them had been aliens, it was also believed that they represented "our" women. Once captured, the woman disappeared from her own community, was brutally treated, whipped with rawhide, and became, as the House Report said, practically a slave in the true sense of the word.

To meet this assumed situation, the White Slave Traffic Act made it a crime to transport a woman "for the purpose of prostitution, or debauchery or for any other immoral purpose."

While Representative Richardson said that the bill was "impractical, vague, indistinct and indefinite in every respect," the debates show that Congress had in mind some very fundamental issues. On one side were those who were in favor of home rule or the powers of the states. . . . On the other side were those who argued that "public health and public morals appeal to us." They were the ones who said, "The proposed legislation is constitutional, and it is related to moral considerations of the most compelling force. If it were not true that our penal legislation were related to moral questions, and moral considerations, then the whole fabric of that legislation would lose its power to command the approbation of the country." They were careful to insist, however, that "the sections proposed do not amount to an interference with the police powers of the States."

In a way the bill must surely have been about the "white-slave traffic." Congressman Mann at the end of the debate emphasized the subject matter by declaring: "Congress would be derelict in its duty if it did not exercise [power], because all the horrors which have been urged, either truthfully or fancifully, against the black-slave traffic pale into insignificance as compared with the horrors of the so-called 'white-slave traffic.'" Congressman Peters said: "The considerations which prompt the support of this bill are so widespread and its objects are so well understood and meet with such universal approval that no explanation or repetition of them need be made to this House. The bill aims to aid in the suppression of the white-slave traffic. . . . " The majority report of the House Committee had defined the white-slave trade as the "business of securing white women and girls and selling them outright or exploiting them for immoral purposes." It had stressed the international character of the trade and the large earnings involved.

Yet while it was said that "the traffic at which this bill strikes is admitted to be abhorrent to all men," and "the time will never arrive when there will be a change of sentiment with respect to its infamy and depravity," there was confusion both as to the facts and as to the legislation proposed.

For example, Congressman Richardson said he knew of the complaint about the traffic but "it may be that there is a good deal of exaggeration about it." Many of the situations described in the House Report had to do with conditions in Illinois. But the "law in Illinois has been strengthened and there had been many prosecutions under it." Congressman Adamson noted that "the Chairman of the Committee on Immigration and Naturalization . . . stated . . . that the white slave traffic had practically been stamped out of our large cities." An examination of the instances cited, Adamson thought, would show that they could be handled under existing laws. And despite the descriptions of immorality, the truth was that society was getting better and "we are vastly better morally than the rest of the world." On the other hand, Congressman Russell, as his contribution on the facts, told the House the story of a Negro who was supposed to have purchased his third white wife "out of a group of twenty-five that were offered for sale in Chicago."

Whatever the evil, presumably the legislation was molded to cure it. The Act speaks of "prostitution," "debauchery," and of "other immoral practice" or "purpose." So far as prostitution was concerned, the report of the House Committee had said that "the bill reported does not endeavor to regulate, prohibit or punish prostitution or the keeping of places where prostitution is indulged in." Congressman Adamson noted that the purpose of the bill "is not to stamp out prostitution, nor do its advocates so contend." He realized that many good men and women, and some good congressmen, thought the purpose was to stamp out prostitution and immorality, but this was an error. But the House Report clearly said that the bill reaches the transportation of women "for the purpose of prostitution." And Representative Peters joined the three elements of white-slave traffic, transportation, and prostitution together in his statement that "the bill aims to aid in the suppression of the white slave traffic by making it a felony to purchase interstate transportation for any woman going to a place for purposes of prostitution. . . . "

Except for the charge that the provisions of the bill were "liable to furnish boundless opportunity to hold up and blackmail and make unnecessary trouble without the corresponding benefit to society," there seemed to be general agreement that the women involved were victims. The women were under control of keepers and were unable to communicate with the outside world without

permission. It was said that the evidence "shows that many victims of this traffic, have been coerced into leading lives of shame, by the use of force, deceit, fraud and every variety of trickery. In many instances they are most unwilling victims, who are literally compelled to practice immorality, and are held to its pursuit by means of violence and restraint." The House Report said that in many cases the women "are practically slaves in the true sense of the word." . . .

There was hardly any discussion of the meaning of "other immoral practice" or "purpose." It was known to some that in 1908, in the *Bitty* case, the Supreme Court had construed similar language in a related statute which dealt with alien women imported for the purpose of prostitution or "any other immoral purpose" to include the importation of a woman for the purpose of concubinage.

The Mann Act was passed after there had been many extensive governmental investigations. Yet there was no common understanding of the facts, and whatever understanding seems to have been achieved concerning the white-slave trade in retrospect seems incorrectly based. The words used were broad and ambiguous. There were three key phrases: "prostitution," "debauchery," and "for any other immoral purpose." The Act was now ready for interpretation.

NOTE

1. *United States v. American Trucking Ass'n,* 310 U.S. 534, 542 (1940).

REVIEW AND DISCUSSION QUESTIONS

1. Explain why Levi thinks the intent of the legislature is important to interpreters of statutes.
2. Why does he think statutory interpretation differs from interpreting case law?
3. What motivated Congress to pass the Mann Act?

Cases Interpreting the Mann Act

Caminetti v. United States

Passed in 1909, the Mann Act (which Congress said should be "known and referred to as the 'White-Slave Traffic Act'") made it a crime to transport across a state line "any woman or girl for the purpose of prostitution or debauchery, or for any other immoral purpose, or with the intent and purpose to induce, entice or compel such woman or girl to become a prostitute, or to give herself up to debauchery, or to engage in any other immoral practice." The Mann Act was upheld by the U.S. Supreme Court in *Hoke v. United States* (1913) under the power given to Congress to regulate interstate commerce.

In this case, *Caminetti v. United States,* four men had been convicted under the act. Two of them had taken their mistresses across a state line; two others coerced an eighteen-year-old girl to go from Oklahoma to Kansas, intending that she engaged in prostitution. In this selection, the Supreme Court focuses on the first two, asking if by taking their mistresses into another state they violated the act.

242 U.S. 470 (1917).

Mr. Justice Day delivered the opinion of the Court: It is contended that the Act of Congress is intended to reach only "commercialized vice" or the traffic in women for gain, and that the conduct for which the several petitioners were indicted and convicted, however reprehensible in morals, is not within the purview of the statute when properly construed in the light of its history and the purposes intended to be accomplished by its enactment. In none of the cases was it charged or proved that the transportation was for gain or for the purpose of furnishing women for prostitution for hire, and it is insisted that, such being the case, the acts charged and proved, upon which conviction was had, do not come within the statute.

It is elementary that the meaning of a statute must, in the first instance, be sought in the language in which the act is framed, and if that is plain, and if the law is within the constitutional authority of the law-making body which passed it, the sole function of the courts is to enforce it according to its terms.

Where the language is plain and admits of no more than one meaning the duty of interpretation does not arise and the rules which are to aid doubtful meanings need no discussion. There is no ambiguity in the terms of this act. It is specifically made an offense to knowingly transport or cause to be transported, etc., in interstate commerce, any woman or girl for the purpose of prostitution or debauchery, or for "any other immoral purpose," or with the intent or purpose, to induce any such woman or girl to become a prostitute or to give herself up to debauchery, or to engage in any other immoral practice.

Statutory words are uniformly presumed, unless the contrary appears, to be used in their ordinary and usual sense. . . . To cause a woman or girl to be transported for the purposes of debauchery, and for an immoral purpose, to-wit, becoming a concubine or mistress. . . . would seem by the very statement of the facts to embrace transportation for purposes denounced by the act, and therefore fairly within its meaning.

While such immoral purpose would be more culpable in morals and attributed to baser motives if accompanied with the expectation of pecuniary gain, such considerations do not prevent the lesser offense against morals of furnishing transportation in order that a woman may be debauched, or become a mistress or a concubine from being the execution of purposes within the meaning of this law. To say the contrary would shock the common understanding of what constitutes an immoral purpose. . . .

. . . [I]t may be observed that while the title of an act cannot overcome the meaning of plain and unambiguous words used in its body, the title of this act embraces the regulation of interstate commerce "by prohibiting the transportation therein for immoral purposes of women and girls, and for other purposes." It is true that §8 of the act provides that it shall be known and referred to as the "White-slave traffic Act," and the report accompanying the introduction of the same into the House of Representatives set forth the fact that a material portion of the legislation suggested was to meet conditions which had arisen in the past few years, and that the legislation was needed to put a stop to a villainous interstate and international traffic in women and girls. Still, the name given to an act by way of designation or description, or the report which accompanies, cannot change the plain import of its words. . . .

Reports to Congress accompanying the introduction of proposed laws may aid the courts in reaching the true meaning of the legislature in cases of doubtful interpretation. But, as we have already said, and it has been so often affirmed as to become a recognized rule, when words are free from doubt they must be taken as the final expression of the legislative intent, and are not to be added to or subtracted from by considerations drawn from titles or designating names or reports accompanying their introduction, or from any extraneous source. In other words, the language being plain, and not leading to absurd or wholly impracticable consequences, it is the sole evidence of the ultimate legislative intent. . . .

The judgment in each of the cases is

Affirmed.

Mr. Justice McKenna, with whom concurred the Chief Justice [White] and Mr. Justice Clarke, dissenting: Undoubtedly in the investigation of the meaning of a statute we resort first to its words, and when clear they are decisive. The principle has attractive and seemingly disposing

simplicity, but that it is not easy of application or, at least, encounters other principles, many cases demonstrate. The words of a statute may be uncertain in their signification or in their application. If the words be ambiguous, the problem they present is to be resolved by their definition; the subject-matter and the lexicons become our guides. But here, even, we are not exempt from putting ourselves in the place of the legislators. If the words be clear in meaning but the objects to which they are addressed be uncertain, the problem then is to determine the uncertainty. And for this a realization of conditions that provoked the statute must inform our judgment. . . .

The transportation which is made unlawful is of a woman or girl "to become a prostitute or to give herself up to debauchery, or to engage in any other immoral practice." Our present concern is with the words "any other immoral practice," which, it is asserted, have a special office. The words are clear enough as general descriptions; they fail in particular designation. . . . "Immoral" is a very comprehensive word. It means a dereliction of morals. In such sense it covers every form of vice, every form of conduct that is contrary to good order. It will hardly be contended that in this sweeping sense it is used in the statute. But if not used in such sense, to what is it limited and by what limited? If it be admitted that it is limited at all, that ends the imperative effect assigned to it in the opinion of the court. But not insisting quite on that, we ask again. By what is it limited? By its context, necessarily, and the purpose of the statute.

For the context I must refer to the statute; of the purpose of the statute Congress itself has given us illumination. It devotes a section to the declaration that the "Act shall be known and referred to as the 'White-slave traffic Act.'" And its prominence gives it prevalence in the construction of the statute. It cannot be pushed aside or subordinated by indefinite words in other sentences, limited even there by the context. . . . The designation "White-slave traffic" has the sufficiency of an axiom. If apprehended, there is no uncertainty as to the conduct it describes. It is commercialized vice, immoralities having a mercenary purpose, and this is confirmed by other circumstances.

The author of the bill was Mr. Mann, and in reporting it from the House Committee on Interstate and Foreign Commerce he declared for the Committee that it was not the purpose of the bill to interfere with or usurp in any way the police power of the States, and further that it was not the intention of the bill to regulate prostitution or the places where prostitution or immorality was practiced, which were said to be matters wholly within the power of the States and over which the federal government had no jurisdiction. And further explaining the bill, it was said that the sections of the act had been "so drawn that they are limited to cases in which there is the act of transportation in interstate commerce of women for purposes of prostitution." And again:

> . . . The legislation is needed to put a stop to a villainous interstate and international traffic in women and girls. . . . It does not attempt to regulate the practice of voluntary prostitution, but aims solely to prevent panderers and procurers from compelling thousands of women and girls against their will and desire to enter and continue in a life of prostitution. . . .

In other words, it is vice as a business at which the law is directed, using interstate commerce as a facility to procure or distribute its victims. . . .

This being the purpose, the words of the statute should be construed to execute it, and they may be so construed even if their literal meaning be otherwise.

REVIEW AND DISCUSSION QUESTIONS

1. Describe the reasoning of Justice Day in upholding the conviction.
2. How does Justice Day think the Court should decide the meaning of a statute?
3. Why does Justice McKenna dispute Day's argument?
4. Which position, the majority or the dissent, is more sound? Explain.
5. In what ways, if any, is Levi's discussion of the legislative history of the Mann Act relevant to this case?

Mortensen v. United States

Mr. and Mrs. Mortensen ran a house of prostitution in Nebraska. Unfortunately for them, however, they took two of their employees on vacation to Salt Lake City and Yellowstone Park. The ladies resumed their work after returning to Nebraska, again in the employ of the Mortensens. The Mortensens were convicted under the Mann Act, though there was no indication the women had been involved in prostitution while on vacation or that the Mortensens had coerced the women to return to Nebraska.

Mr. JUSTICE MURPHY delivered the opinion of the Court: The penalties of §2 of the Act are directed at those who knowingly transport in interstate commerce "any woman or girl for the purpose of prostitution or debauchery, or for any other immoral purpose, or with the intent and purpose to induce, entice, or compel such woman or girl to become a prostitute or to give herself up to debauchery, or to engage in any other immoral practice." The statute thus aims to penalize only those who use interstate commerce with a view toward accomplishing the unlawful purposes. To constitute a violation of the Act, it is essential that the interstate transportation have for its object or be the means of effecting or facilitating the proscribed activities. An intention that the women or girls shall engage in the conduct outlawed by §2 must be found to exist before the conclusion of the interstate journey and must be the dominant motive of such interstate movement. And the transportation must be designed to bring about such result. Without that necessary intention and motivation, immoral conduct during or following the journal is insufficient to subject the transporter to the penalties of the act. . . .

It may be assumed that petitioners anticipated that the two girls would resume their activities as prostitutes upon their return to Grand Island. But we do not think it is fair or permissible under the evidence adduced to infer that this interstate vacation trip, or any part of it, was undertaken by petitioners for the purpose of, or as a means of effecting or facilitating, such activities. The sole purpose of the journey from beginning to end was to provide innocent recreation and a holiday for petitioners and the two girls. . . . What Congress has outlawed by Mann Act, however, is the use of interstate commerce as a calculated means for effectuating sexual immorality. In ordinary speech an interstate trip undertaken for an innocent vacation purpose constitutes the use of interstate commerce for that innocent purpose. Such a trip does not lose that meaning when viewed in light of a criminal statute outlawing interstate trips for immoral purposes.

The fact that the two girls actually resumed their immoral practices after their return to Grand Island does not, standing alone, operate to inject a retroactive illegal purpose into the return trip to Grand Island. Nor does it justify an arbitrary splitting of the round trip into two parts so as to permit an inference that the purpose of the drive to Salt Lake City was innocent while the purpose of the homeward journey to Grand Island was criminal. The return journey under the circumstances of this case cannot be considered apart from its integral relation with the innocent round trip as a whole. There is no evidence of any change in the purpose of the trip during its course. If innocent when it began it remained so until it ended. . . .

To punish those who transport inmates of a house of prostitution on an innocent vacation trip in no way related to the practice of their commercial vice is consistent neither with the purpose nor with the language of the Act. Congress was attempting primarily to eliminate the "white-slave" business. . . . To accomplish its purpose the statute enumerates the prohibited acts in broad language capable of application beyond that intended by the legislative framers. But even such broad language is conditioned upon the use of interstate transportation for the purpose of, or as a means of effecting or facilitating, the commission of the illegal acts. Here the interstate round trip had no such

322 U.S. 369 (1944).

purpose and was in no way related to the subsequent immoralities in Grand Island. . . .

The judgment of the court below is

Reversed.

Mr. Chief Justice Stone: Mr. Justice Black, Mr. Justice Reed, Mr. Justice Douglas and I think the judgment should be affirmed: Courts have no more concern with the policy and wisdom of the Mann Act than of the Labor Relations Act or any other which Congress may constitutionally adopt. Those are matters for Congress to determine, not the courts. . . .

The fact that petitioners, who were engaged in an established business of operating a house of prostitution in Nebraska, took some of its women inmates on a transient and innocent vacation trip to other states, is in no way incompatible with the conclusion that petitioners, in bringing them back to Nebraska, purposed and intended that they should resume there the practice of commercial vice, which in fact they did promptly resume in petitioners' establishment. The record is without evidence that they engaged or intended to engage in any other activities in Nebraska, or that anything other than the practice of their profession was the object of their return.

REVIEW AND DISCUSSION QUESTIONS

1. Describe the reasons Justice Murphy gives for reversing the conviction.
2. On what basis do the dissenters claim the convictions should be upheld? Which side do you think is correct?
3. Discuss the relevance for this case of the legislative history of the Mann Act described by Edward Levi.
4. Is this decision compatible with *Caminetti?* Explain.

Cleveland v. United States

This case raised still further questions about the reach of the Mann Act. Six members of the Mormon Church were convicted for violating the act by transporting their wives across state lines. The reason for their conviction was that the Mormons practiced polygamy.[1]

Mr. Justice Douglas delivered the opinion of the Court: The Act makes an offense the transportation in interstate commerce of "any woman or girl for the purpose of prostitution or debauchery, or for any other immoral purpose." The decision turns on the meaning of the latter phrase, "for any other immoral purpose." . . .

It is argued that the *Caminetti* decision gave too wide a sweep to the Act; that the Act was designed to cover only the white slave business and related vices; that it was not designed to cover voluntary actions bereft of sex commercialism; and that in any event it should not be construed to embrace polygamy which is a form of marriage and, unlike prostitution or debauchery or the concubinage involved in the *Caminetti* case, has as its object parenthood and the creation and maintenance of family life. . . .

While *Mortensen v. United States* [1944] rightly indicated that the Act was aimed "primarily" at the use of interstate commerce for the conduct of the white slave business, we find no indication that a profit motive is a *sine qua non* to its application. Prostitution, to be sure, normally suggests sexual relations for hire. But debauchery has no such implied limitation. In common understanding the

[1] 329 U.S. 14 (1946).

indulgence which that term suggests may be motivated solely by lust. And so we start with words which by their natural import embrace more than commercialized sex. What follows is "any other immoral purpose." Under the *ejusdem generis* [of the same class or kind] rule of construction the general words are confined to the class and may not be used to enlarge it. But we could not give the words a faithful interpretation if we confined them more narrowly than the class of which they are a part.

That was the view taken by the Court in [earlier] cases. We do not stop to reexamine the *Caminetti* case to determine whether the Act was properly applied to the facts there presented. But we adhere to its holding . . . that the Act, while primarily aimed at the use of interstate commerce for the purposes of commercialized sex, is not restricted to that end.

We conclude, moreover, that polygamous practices are not excluded from the Act. They have long been outlawed in our society. As stated in *Reynolds v. United States* [1879]:

> Polygamy has always been odious among the northern and western nations of Europe, and, until the establishment of the Mormon Church, was almost exclusively a feature of the life of Asiatic and of African people. At common law, the second marriage was always void (2 Kent, Com. 79), and from the earliest history of England polygamy has been treated as an offence against society. . . .

. . . Polygamy is a practice with far more pervasive influences in society than the casual, isolated transgressions involved in the *Caminetti* case. The establishment or maintenance of polygamous households is a notorious example of promiscuity. . . . We could conclude that Congress excluded these practices from the Act only if it were clear that the Act is confined to commercialized sexual vice. Since we cannot say it is, we see no way by which the present transgressions can be excluded. These polygamous practices have long been branded as immoral in the law. Though they have different ramifications, they are in the same genus as the other immoral practices covered by the Act. . . .

It is also urged that the requisite criminal intent was lacking since petitioners were motivated by a religious belief. That defense claims too much. If upheld, it would place beyond the law any act done under claim of religious sanction. But it has long been held that the fact that polygamy is supported by a religious creed affords no defense in a prosecution for bigamy. *Reynolds v. United States.* Whether an act is immoral within the meaning of the statute is not to be determined by the accused's concepts of morality. Congress has provided the standard. The offense is complete if the accused intended to perform, and did in fact perform, the act which the statute condemns. . . .

Affirmed.

Mr. Justice Murphy, dissenting: Today another unfortunate chapter is added to the troubled history of the White Slave Traffic Act. It is a chapter written in terms that misapply the statutory language and that disregard the intention of the legislative framer. It results in the imprisonment of individuals whose actions have none of the earmarks of white slavery, whatever else may be said of their conduct. . . .

It is not my purpose to defend the practice of polygamy or to claim that it is morally the equivalent of monogamy. But it is essential to understand what it is, as well as what it is not. Only in that way can we intelligently decide whether it falls within the same genus as prostitution or debauchery.

There are four fundamental forms of marriage: (1) monogamy; (2) polygyny, or one man with several wives; (3) polyandry, or one woman with several husbands; and (4) group marriage. The term "polygamy" covers both polygyny and polyandry. Thus we are dealing here with polygyny, one of the basic forms of marriage. Historically, its use has far exceeded that of any other form. It was quite common among ancient civilizations and was referred to many times by the writers of the Old Testament; even today it is to be found frequently among certain pagan and non-Christian peoples of the world. We must recognize, then, that polygyny, like other forms of marriage, is basically a cultural institution rooted deeply in the religious beliefs and social mores of those societies in which it appears. It is equally true that the beliefs and mores of the dominant culture of the contemporary world condemn the practice as immoral and substitute monogamy in its place. To those beliefs and mores I subscribe, but that does not alter the

fact that polygyny is a form of marriage built upon a set of social and moral principles. . . .

The Court states that polygamy is "a notorious example of promiscuity." The important fact, however, is that, despite the differences that may exist between polygamy and monogamy, such differences do not place polygamy in the same category as prostitution or debauchery. When we use those terms we are speaking of acts of an entirely different nature, having no relation whatever to the various forms of marriage. It takes no elaboration here to point out that marriage, even when it occurs in a form of which we disapprove, is not to be compared with prostitution or debauchery or other immoralities of that character.

The Court's failure to recognize this vital distinction and its insistence that polygyny is "in the same genus" as prostitution and debauchery do violence to the anthropological factors involved. Even etymologically, the words "polygyny" and "polygamy" are quite distinct from "prostitution," "debauchery," and words of that ilk. There is thus no basis in fact for including polygyny within the phrase "any other immoral purpose" as used in this statute.

REVIEW AND DISCUSSION QUESTIONS

1. Why does Justice Douglas think polygamous practices are included in the act?
2. How is the precedent found in *Reynolds v. United States* relevant to the case, according to Justice Douglas?
3. Summarize briefly the reasons Justice Murphy dissents.
4. Explain what role (if any) is played in the opinions by the justices' attitudes toward polygamy.
5. Trace the development of the Court's thinking about the meaning of the Mann Act from *Caminetti* to this case. Is there a single, consistent pattern here?
6. Which of the three Mann Act cases is most incompatible with the legislative history Levi describes?
7. Why is it important for courts to rely on the intent of the legislature when they interpret a statute?

Interpretation

Tony Honoré

Interpretation, as Tony Honoré points out, is of great importance to the law. But what, precisely, is interpretation? Beginning with brief remarks about how we interpret one another's meaning in ordinary life, Honoré goes on to discuss the formal nature of legal interpretation and the strengths of different approaches to interpretation. He provides an example of interpretation involving drinking and driving, and he explains why it is the case that "words cannot mean just anything." He then discusses the role of legislative intention in interpretation. Tony Honoré was professor of philosophy at Oxford.

Interpretation is one of the main concerns of lawyers.

Originally, lawyers were people who specialized in drafting and interpreting documents, rather than arguing cases in court or judging disputes. A person drafting a document must try to see in advance how the law, treaty, contract or will he is drafting will be interpreted, so that his text covers the areas he wants it to cover. Arguments in court are often about the interpretation of a text, and judges have to decide what the right interpretation is. So, one way and another, interpretation is a key part of legal practice.

How documents should be interpreted is disputed. Suppose a by-law forbids vehicles in a park.

Does this apply to bicycles, motorcycles or powered lawn-mowers? More important, how do you set about deciding? Do you go by the dictionary meaning of the word "vehicle"? Do you ask what purpose banning vehicles from the park is meant to serve? Do you ask what hardship to the public would be caused if, for example, bicycles or baby-carriages were included in the ban? Or how inconvenient it would be if powered lawn-mowers were excluded?

Some lawyers give priority to the letter of the text, others to its spirit or purpose, others again to the likely result of interpreting it this way or that. This chapter is mainly concerned with the clash between these different approaches.

WHAT IS INTERPRETATION?

First of all, what is interpretation? To interpret what someone says is to attach a meaning to their remarks. When the meaning is at once clear the remark does not have to be interpreted, but if it is not clear, or not immediately clear, it does. The interpreter then has to choose between two or more possible ways of understanding what was said.

Wanting to travel to Manchester, you ask me to get you a time-table. I am not sure whether you meant a coach time-table or a railway time-table. I may be able to ask you which you meant. But if I cannot, I have to make up my mind, perhaps on the basis of how you usually travel, or what the easiest way of getting to Manchester is.

Interpreting a request of this sort is not just interpreting the words "please get me a time-table," which are straightforward enough, but interpreting them in the context of a purpose, a wish to travel to Manchester. Interpreting texts, which is so central a part of a lawyer's work, is in some ways similar. A lawyer interprets a text as part of a statute, contract, will, treaty, regulation or whatever. But there are differences. When I have to interpret your request for a time-table I am free to go about it in any way I choose. A lawyer's interpretation of a text is a more formal process.

Legal Interpretation Is Formal

It is formal, first, because the texts a lawyer interprets are in writing. Not only are they in writing but they have authority. If they are statutes they are part of state law. If they are contracts or treaties they bind the parties to the contract or treaty. If they are wills they bind whoever is dealing with the property of the person who has died. The interpretation chosen will make a difference to someone's rights and duties.

There is another way in which the interpretation of a legal text is formal. When it is disputed, there is such a thing as an official interpretation of the text. Judges provide this when they try cases or hear appeals. Ministers and civil servants, for example tax officials, also issue official interpretations of statutes, though their interpretations have in the end to give way to those of judges if there is a difference of opinion between the two.

Another way in which the interpretation of legal texts is unlike interpreting a request from a friend is that the evidence on which the interpreter of legal texts has to come to a decision is limited. If I am not sure what my friend meant by his request my best move is to ask him. But the interpreter of a legal text cannot solve the problem by going back to the author of the text and asking what he meant. For one thing, legal texts often have no single author. They are enacted by legislators, hundreds in number, or agreed by the parties to a contract or treaty, of whom there may be many.

Legislators do not have time to explain what they meant by the statutes they have passed. Even if they did, and all agreed, it might not be a good idea to ask them, because they would be tempted to explain the law with the benefit of hindsight. They would be tempted to put themselves in a better light by saying that it meant something that they would not have said it meant at the time it was passed. This would not be fair to those who have to obey the law. They are entitled to be judged by what in the view of an impartial person the law meant at the time they had to obey it.

The interpreter of a will obviously cannot consult the testator about what it meant, as he is dead. It might seem that the interpreter of a contract or treaty could consult the parties to the treaty or contract. But if a dispute has arisen it is often because the parties disagree about how it should be interpreted. A neutral interpretation is called for. So to ask the parties is not a real option. In general,

problems of interpreting legal texts cannot be solved by going back to the authors of the text.

Another way in which the evidence available to the interpreter of a text is limited is that the text has to speak for itself. Earlier drafts of a statute, will, contract or treaty are superseded by the final version. The authors of the text are committed to the final version, and this replaces whatever discussions or negotiations went on beforehand. Even what the authors said they meant at the time the text was agreed upon is treated with reserve. For instance, some systems of law do not accept that in interpreting the text of a statute the interpreter can go by what the Minister said it meant when he spoke in the legislature.

How far one presses the idea that a text should speak for itself is a matter of dispute, but at its core there lies a sound idea. This is that a person who agrees that a certain form of words is to be binding should be willing for the words to be taken in their ordinary sense. Otherwise he should have insisted on a different form of words or said at the time that he attached a special meaning to them, not the ordinary meaning.

Attaching a special meaning is in fact quite a common practice. Statutes and contracts often have clauses that lay down the meaning to be given to the words they use. For instance they may say that "owner" includes the occupier of a house. In that case a tenant who does not own the house is treated for purposes of the statute (but not otherwise) as if he did. Behind this practice lies the view that in legal texts the words should be taken in their ordinary sense unless the author of the text lays down a different sense.

WHAT IS THE BEST APPROACH TO INTERPRETATION?

The Textual Approach

The idea that a text must speak for itself embodies one approach to interpretation, which I shall call the textual approach. It holds that a text must be understood in its ordinary sense. If vehicles are excluded from the park that means that what are ordinarily spoken of as vehicles (cars, coaches, trucks) are excluded. This may be inconvenient if it means, for example, that some disabled people cannot use the park because they can only get there by a motorized vehicle. But that has to be accepted.

Sometimes of course a word or phrase has no ordinary sense, because it is not one used by ordinary people. When technical terms, legal or scientific, are used, what corresponds to the ordinary meaning of the term is its technical meaning. The technical meaning is what it ordinarily means to legal or scientific experts.

But those who favour the textual approach admit that there are exceptional cases where it is unsatisfactory. When it would produce an absurd result to read the words in their ordinary sense, they can be interpreted in a different sense, one that does not produce an absurd result.

What amounts to an absurd result is of course debatable; what is absurd to one judge is merely unusual to another. Would it be absurd if the meaning given to "vehicle" resulted in excluding some disabled people from the park? A text that is self-contradictory or unworkable in practice is clearly absurd, but how far one should treat a meaning that some people think unreasonable as absurd is not clear.

However we understand absurdity, the textual approach gives priority to the language used in the text in its ordinary sense over other evidence of the author's intention. The textual approach is sometimes attacked by critics, who call it "literalism," going by the letter. But what is the point of putting a statute, contract, treaty, or will into words unless those words are to be treated as binding?

The Purposive Approach

A different approach to interpreting texts is the purposive approach. According to this, priority in interpreting texts should be given to the purpose of the statute, contract, treaty, will, or regulation. Account should also be taken of the general aims of the legal system, such as justice, security, and efficiency. This is specially important in interpreting statutes. A law (say the law forbidding vehicles in the park) can have special purposes of its own (say to ensure peace and quiet for those who use the park). But more general purposes have also to be taken into account (for instance that public places should be

kept tidy, which may require powered lawn-mowers to be admitted even if they make a noise).

If these special or general purposes point to an interpretation different from the ordinary meaning of the words, supporters of the purposive approach say that the "purposive" meaning should be adopted in preference to the ordinary meaning.

In looking for the purpose of a document there are, however, some limits on how far afield one can range. Legal systems vary a bit in the limits they set. When a statute is interpreted the interpreter can certainly look at the whole statute and any earlier laws on the subject. He can take account of what the previous state of the law was and what were thought to be its defects. Supporters of the purposive approach go further, and say that the interpreter should be able to take account of reports and proposals of official commissions and other materials prepared with a view to passing a law or making a treaty. Many systems of law do in fact allow these materials to be considered.

A "Drinking and Driving" Example

Which approach is better, the textual or the purposive approach? I take an example (altering it a bit) from a law in Britain that has since been replaced. Suppose a statute lays down that a policeman may require someone who is driving a vehicle to take a blood test. If the blood test shows that he has more than a certain amount of alcohol in the blood, he may be charged with the offence of drinking and driving.

What does "driving" mean in this statute? Like most words, "driving" has a range of meanings. But in its ordinary meaning, many people would say, a person is driving a vehicle when he is at the controls and the vehicle is moving. When the vehicle stops he stops driving. But if we interpret driving in this way, we make the statute absurd. How could a policeman tell a driver to take a blood test while the vehicle was moving?

So, even on the textual approach, we ought to ask if "driving" can be understood in another way. If the statute is to have some point, the driver must still be driving after he has stopped the vehicle because the policeman has told him to. Can driving be understood in this wider sense? It probably can, since from one point of view the driver of a vehicle is driving from the moment he sets off until he gets to his destination, even if he stops on the way for a meal. If asked during the meal what he was doing he would say he was "driving to Manchester." So does "driving" in the statute mean simply that the driver is still on the way to his destination?

From the words alone this is a possible interpretation. But ought not the purpose of the law to be taken into account when we interpret "driving?" The point of the law is surely to make the person driving liable to take a blood test; and the purpose of this is to reduce the danger that the motorist will drive with more than a certain amount of alcohol in his blood. Would it fit that purpose to give a blood test to someone having a meal who has not been driving for the last half-hour and may not set off again for some time? Would it be fair to him?

Perhaps, taking the purpose of the law into account, a driver is "driving" only if he has already started and means to go on with the journey the moment he can. It is only then that, if the alcohol in his blood is over the limit, he will be a danger on the road. In that case he may be driving though he is temporarily at a halt in a traffic jam, or because a policeman has stopped him. If the driver has stopped for a meal he is not driving for the purpose of the blood test, but the driver who has got out of his car in a traffic jam to buy a paper is still driving.

On the textual approach the ordinary meaning of "driving" (that the vehicle must be in motion) leads to an absurdity, so that other ways of understanding "driving" can be considered. On the purposive approach the ordinary meaning of "driving" has no special priority and a motorist can be said to be driving when it would promote the purpose of the drinking and driving law to give him a blood test. This might suggest a very wide interpretation of driving, one that would include the stage when he is having a meal on the way.

But the purpose of reducing drinking and driving is not the only purpose that the interpreter has to take into account. There are also general aims of the legal system. One of these is to be fair to people who may be charged with a crime. They should be able to know roughly when they are in danger of being tested and prosecuted. This brings us back by a roundabout route to the ordinary meaning of "driving." When a statute uses a term that is part of ordinary language, those affected by it are likely to

understand it in its ordinary sense. So the purposes of the legal system itself require close attention to be paid to the words used in the text.

Words Cannot Mean Just Anything

Though words have a range of meanings, they cannot mean just anything. For instance, there are some things that "driving" could not mean. A driver whose vehicle is in the garage and who has not yet set off, though he intends to leave shortly, and who is in the meantime having a few drinks, is not driving. It is true that, if he were held to be driving, the purpose of the law against drinking and driving would be satisfied. But this cannot justify the interpreter in stretching the meaning of driving to cover the driver whose vehicle is still in the garage.

It is very rare for a judge to decide that a text means something that it could not mean in ordinary or technical language. On the few occasions when a judge does this, he does it because he thinks that an obvious mistake has been made by the author of the text, and that he has a duty to correct it.

There is little doubt that "or" cannot mean "and." Elephants and hyenas means both species, but elephants or hyenas means one of the two. So if a statute lays down that no one may hunt elephants and hyenas without a licence, the draftsman has probably made a mistake. As it stands, the law seems to say that one may hunt either elephants or hyenas without a licence but not both. It is unlikely that this is what the legislators meant. Most hunters will be hunting one or the other, not both, and the prohibition will then have very little effect.

So if a judge had to interpret this statute he would probably interpret "and" as "or," and say that "elephants and hyenas" should read as if it said "elephants or hyenas." That seems to go against the textual approach, by taking a word to mean what it cannot mean. It seems as if the judge is opting for the purposive approach, in order to make the statute workable. A different way of putting it, more consistent with the textual approach, is that the judge has the power to correct an obvious mistake by substituting "or" for "and." In that case he is not interpreting the text but correcting it, something that can obviously be allowed only within very narrow limits.

Intention

The textual approach and the purposive approach are often put forward as different ways of getting at the intention of the author of the text. That is not quite right, because the textual approach opts for what the words ordinarily mean, and the author of the text may not have meant to use the actual words he did (take the elephants and hyenas example).

But assuming the author of the text did mean to use the words he used, there is sometimes a contrast between what he meant by the words used and what he meant to achieve by using them. These are two different things, though they can both be called his "intention." Does it help to think of interpretation as a search for the intention of the legislators, the parties to the contract, etc.?

Individuals have both sorts of intention, bodies like legislatures and associations can have both too. But often the members of these bodies do not study the texts that are produced in their name at all carefully. They can say roughly what they want to achieve by passing a law. They may for example be against drinking and driving and in favour of animal conservation. But they probably have not thought about what is meant by "driving" in the blood testing statute or "elephants and hyenas" in the hunting statute.

All the same there is good reason, I think, to say that the interpreter should try to discover the intention of the legislature or the parties to a contract or treaty. A statute, contract or treaty is a compromise between different views. Perhaps no member of the legislature, and no party to the contract or treaty, would themselves have chosen the text that was finally agreed, if it depended on them alone. The point of speaking of the intention of the legislature or the contracting parties is not that any particular person's views should govern the interpretation of the text. It is rather that the interpreter should treat the text as if it represented the views of a single individual, and make it as coherent as the words permit.

This is a fiction, but it expresses a sound policy. The interpreter should not treat the views of individual members of a large body as if they were the views of the whole body. There may have been no agreed view about what the text meant. In the same way, the views of one party to a contract about what

the contract means should be not be treated as if they were the views of both parties. The interpreter should aim at a neutral, impartial reading of the text and its purposes.

Whether the textual or the purposive approach to interpretation is preferred, every text has to be seen against the background of the society and the legal system of which it forms part.

REVIEW AND DISCUSSION QUESTIONS

1. In what senses does Honoré think that legal interpretation is more "formal" than the interpretation of a railroad time-table?
2. Describe the textual and the purposive approaches to interpretation.
3. Using the statute involving drinking and driving, discuss Honoré's position on the advantages of the different approaches to interpretation.
4. Why does Honoré think using intentions to interpret is a "fiction"? Why does he nonetheless think it is a useful fiction?

Who Is a Drunk Driver?

Kentucky v. Whitt

The issue in this case involves the interpretation of a statute making it a crime to drive a vehicle while drunk.

It was reported in the *St. Petersburg Times* (October 18, 1992) that Mark A. Whitt had been charged with drunk driving. In 1991 Kentucky had amended its laws to include those who drive "non-motorized vehicles." The facts of the case were not in dispute. Mr. Whitt did not deny that he was drunk when he rode his horse Mabel home; indeed he admitted that he had done the same thing before, after drinking at a club. When told that he had been charged with drunk driving, however, Mr. Whitt reportedly objected saying that his beloved Mabel "has a mind of her own" adding that "I don't think a vehicle has a mind. I've even passed out in the saddle before. She knows the way home." Mr. Whitt faces a fine of up to $100. Police said that this was the first case involving a horse.

The law under which he was charged states that "no person under the influence of intoxicating beverages or any substance that may impair one's driving abilities shall operate a vehicle that is not a motor vehicle anywhere in the state." [K.R.S. Ann., 189.520 (1)] The statute defines a vehicle as "all agencies for the transportation of persons or property over or upon the public highways" (though it exempted construction equipment) and it defined "motor vehicle" as including "all vehicles as defined above which are propelled otherwise than by muscular power." [K.R.S. Ann., 189.010 (18)]

REVIEW AND DISCUSSION QUESTIONS

1. What are the relevant facts in this case?
2. Explain whether you think Mr. Whitt violated the law, the reason why you reached that conclusion, and how the nature of statutory interpretation in general is relevant to your conclusion.

What Is a Vegetable?

Nix v. Hedden

Federal law required people importing vegetables to pay a 10 percent import tax if the vegetable is in its "natural state, or in salt or brine." The law excluded from the tax all fruits. When the government imposed the tax, one importer of tomatoes sued the government claiming he did not owe the tax.

Mr. Justice Gray: At the trial, the plaintiff's counsel, after reading in evidence definitions of the words "fruit" and "vegetables" from *Webster's Dictionary, Worcester's Dictionary* and the *Imperial Dictionary,* called two witnesses, who had been for thirty years in the business of selling fruit and vegetables, and asked them, after hearing these definitions, to say whether these words had "any special meaning in trade or commerce, different from those read."

One of the witnesses answered as follows: "Well, it does not classify all things there, but they are correct as far as they go. It does not take all kinds of fruit or vegetables; it takes a portion of them. I think the words 'fruit' and 'vegetable' have the same meaning in trade to-day that they had on March 1, 1883 [the date the Tariff Act was enacted]. I understand that the term 'fruit' is applied in trade only to such plants or parts of plants as contain the seeds. There are more vegetables than those in the enumeration given in *Webster's Dictionary* under the term 'vegetable,' as 'cabbage, cauliflower, turnips, potatoes, peas, beans, and the like,' probably covered by the words 'and the like.' "

The other witness testified: "I don't think the term 'fruit' or the term 'vegetables' had, in March, 1883, and prior thereto, any special meaning in trade and commerce in this country, different from that which I have read here from the dictionaries."

The plaintiff's counsel then read in evidence from the same dictionaries the definitions of the word "tomato."

The defendant's counsel then read in evidence from *Webster's Dictionary* the definitions of the words "pea," "eggplant," "cucumber," "squash" and "pepper."

The plaintiff then read in evidence from Webster's and Worcester's dictionaries the definitions of "potato," "turnip," "parsnip," "cauliflower," "cabbage," "carrot" and "bean."

No other evidence was offered by either party.

The single question in this case is whether tomatoes, considered as provisions, are to be classed as "vegetables" or as "fruit," within the meaning of the Tariff Act of 1883.

The only witnesses called at the trial testified that neither "vegetables" nor "fruit" had any special meaning in trade or commerce, different from that given in the dictionaries.

The passages cited from the dictionaries define the word "fruit" as the seed of plants, or that part of plants which contains the seed, and especially the juicy, pulpy products of certain plants, covering and containing the seed. These definitions have no tendency to show that tomatoes are "fruit," as distinguished from "vegetables," in common speech, or within the meaning of the Tariff Act.

There being no evidence that the words "fruit" and "vegetables" have acquired any special meaning in trade or commerce, they must receive their ordinary meaning. Of that meaning the court is bound to take judicial notice, as it does in regard to all words in our own tongue; and upon such a question dictionaries are admitted, not as evidence, but only as aids to the memory and understanding of the court.

Botanically speaking, tomatoes are the fruit of a vine, just as are cucumbers, squashes, beans and peas. But in the common language of the people, whether sellers or consumers of provisions, all these are vegetables, which are grown in kitchen

149 U.S. 304 (1893).

gardens, and which, whether eaten cooked or raw, are, like potatoes, carrots, parsnips, turnips, beets, cauliflower, cabbage, celery and lettuce, usually served at dinner in, with or after the soup, fish or meats which constitute the principal part of the repast, and not, like fruits generally, as dessert.

REVIEW AND DISCUSSION QUESTIONS

1. What are the relevant facts in this case?
2. Explain whether you think the government was right to impose the tax, the reason why you reached that conclusion, and how the nature of statutory interpretation in general is relevant to your conclusion.

"Use" of a Firearm

Smith v. United States

Drug trafficking is a serious criminal offense. Moreover, if the defendant "uses" a firearm "during and in relation to . . . [a] drug trafficking crime," federal statute requires an additional five years of incarceration (thirty years if the firearm is a "machinegun" or is fitted with a silencer). The question before the Supreme Court in this case is whether John Angus Smith violated this statute when he agreed to swap his automatic MAC-10 with silencer for two ounces of cocaine. Writing for the majority, Justice O'Connor argues that what Smith did can be described as "use" within the everyday meaning of the term. In his dissent, Justice Scalia argues that to speak of "using a firearm" is to speak of using it as a weapon, which Smith did not do. The arguments of both O'Connor and Scalia raise important issues about intent, meaning, and the principles that should guide statutory interpretation.

Justice O'Connor delivered the opinion of the Court: We decide today whether the exchange of a gun for narcotics constitutes "use" of a firearm "during and in relation to . . . [a] drug trafficking crime" within the meaning of 18 U.S.C. § 924(c)(1). We hold that it does.

I

Petitioner John Angus Smith and his companion went from Tennessee to Florida to buy cocaine; they hoped to resell it at a profit. While in Florida, they met petitioner's acquaintance, Deborah Hoag. Hoag agreed to, and in fact did, purchase cocaine for petitioner. She then accompanied petitioner and his friend to her motel room, where they were joined by a drug dealer. While Hoag listened, petitioner and the dealer discussed petitioner's MAC-10 firearm, which had been modified to operate as an automatic. The MAC-10 apparently is a favorite among criminals. It is small and compact, lightweight, and can be equipped with a silencer. Most important of all, it can be devastating: A fully automatic MAC-10 can fire more than 1,000 rounds per minute. The dealer expressed his interest in becoming the owner of a MAC-10, and petitioner promised that he would discuss selling the gun if his arrangement with another potential buyer fell through.

Unfortunately for petitioner, Hoag had contacts not only with narcotics traffickers but also with law enforcement officials. In fact, she was a confidential informant. Consistent with her post, she informed the Broward County Sheriff's Office of petitioner's activities. The Sheriff's Office responded quickly, sending an undercover officer

508 U.S. 223 (1993).

to Hoag's motel room. Several others were assigned to keep the motel under surveillance. Upon arriving at Hoag's motel room, the undercover officer presented himself to petitioner as a pawnshop dealer. Petitioner, in turn, presented the officer with a proposition: He had an automatic MAC-10 and silencer with which he might be willing to part. Petitioner then pulled the MAC-10 out of a black canvas bag and showed it to the officer. The officer examined the gun and asked petitioner what he wanted for it. Rather than asking for money, however, petitioner asked for drugs. He was willing to trade his MAC-10, he said, for two ounces of cocaine. The officer told petitioner that he was just a pawnshop dealer and did not distribute narcotics. Nonetheless, he indicated that he wanted the MAC-10 and would try to get the cocaine. The officer then left, promising to return within an hour.

Rather than seeking out cocaine as he had promised, the officer returned to the Sheriff's Office to arrange for petitioner's arrest. But petitioner was not content to wait. The officers who were conducting surveillance saw him leave the motel room carrying a gun bag; he then climbed into his van and drove away. The officers reported petitioner's departure and began following him. When law enforcement authorities tried to stop petitioner, he led them on a high-speed chase. Petitioner eventually was apprehended.

II

A grand jury sitting in the District Court for the Southern District of Florida returned an indictment charging petitioner with, among other offenses, two drug trafficking crimes. . . . Most important here, the indictment alleged that petitioner knowingly used the MAC-10 and its silencer during and in relation to a drug trafficking crime. Under 18 U.S.C. § 924(c)(1), a defendant who so uses a firearm must be sentenced to five year's incarceration. And where, as here, the firearm is a "machinegun" or is fitted with a silencer, the sentence is 30 years. . . . The jury convicted petitioner on all counts.

On appeal, petitioner argued that § 924(c)(1)'s penalty for using a firearm during and in relation to a drug trafficking offense covers only situations in which the firearm is used as a weapon. According to petitioner, the provision does not extend to defendants who use a firearm solely as a medium of exchange or for barter. The Court of Appeals for the Eleventh Circuit disagreed. The plain language of the statute, the court explained, imposes no requirement that the firearm be used as a weapon. Instead, any use of "the weapon to facilitate *in any manner* the commission of the offense" suffices. . . .

III

Section 924(c)(1) requires the imposition of specified penalties if the defendant, "during and in relation to any crime of violence or drug trafficking crime[,] uses or carries a firearm." By its terms, the statute requires the prosecution to make two showings. First, the prosecution must demonstrate that the defendant "use[d] or carrie[d] a firearm." Second, it must prove that the use or carrying was "during and in relation to" a "crime of violence or drug trafficking crime."

Petitioner argues that exchanging a firearm for drugs does not constitute "use" of the firearm within the meaning of the statute. He points out that nothing in the record indicates that he fired the MAC-10, threatened anyone with it, or employed it for self-protection. In essence, petitioner argues that he cannot be said to have "use[d]" a firearm unless he used it as a weapon, since that is how firearms most often are used. . . . [W]e confine our discussion to what the parties view as the dispositive issue in this case: whether trading a firearm for drugs can constitute "use" of the firearm within the meaning of § 924(c)(1).

When a word is not defined by statute, we normally construe it in accord with its ordinary or natural meaning. . . . Surely petitioner's treatment of his MAC-10 can be described as "use" within the everyday meaning of that term. Petitioner "used" his MAC-10 in an attempt to obtain drugs by offering to trade it for cocaine. Webster's defines "to use" as "[t]o convert to one's service" or "to employ." Black's Law Dictionary contains a similar definition: "[t]o make use of; to convert to one's service; to employ; to avail oneself of; to utilize; to carry out a purpose or action by means of." Indeed, over 100 years ago we gave the word "use" the same gloss,

indicating that it means "to employ" or "to derive service from." Petitioner's handling of the MAC-10 in this case falls squarely within those definitions. By attempting to trade his MAC-10 for the drugs, he "used" or "employed" it as an item of barter to obtain cocaine; he "derived service" from it because it was going to bring him the very drugs he sought.

In petitioner's view, § 924(c)(1) should require proof not only that the defendant used the firearm, but also that he used it *as a weapon.* But the words "as a weapon" appear nowhere in the statute. Rather, § 924(c)(1)'s language sweeps broadly, punishing any "us[e]" of a firearm, so long as the use is "during and in relation to" a drug trafficking offense.... Had Congress intended the narrow construction petitioner urges, it could have so indicated. It did not, and we decline to introduce that additional requirement on our own.

Language, of course, cannot be interpreted apart from context. The meaning of a word that appears ambiguous if viewed in isolation may become clear when the word is analyzed in light of the terms that surround it. Recognizing this, petitioner and the dissent argue that the word "uses" has a somewhat reduced scope in § 924(c)(1) because it appears alongside the word "firearm." Specifically, they contend that the average person on the street would not think immediately of a guns-for-drugs trade as an example of "us[ing] a firearm." Rather, that phrase normally evokes an image of the most familiar use to which a firearm is put—use as a weapon. Petitioner and the dissent therefore argue that the statute excludes uses where the weapon is not fired or otherwise employed for its destructive capacity.... Indeed, relying on that argument—and without citation to authority—the dissent announces its own, restrictive definition of "use." "To use an instrumentality," the dissent argues, "ordinarily means to use it for its intended purpose."

There is a significant flaw to this argument. It is one thing to say that the ordinary meaning of "uses a firearm" *includes* using a firearm as a weapon, since that is the intended purpose of a firearm and the example of "use" that most immediately comes to mind. But it is quite another to conclude that, as a result, the phrase also *excludes* any other use. Certainly that conclusion does not follow from the phrase "uses . . . a firearm" itself. As the dictionary definitions and experience make clear, one can use a firearm in a number of ways. That one example of "use" is the first to come to mind when the phrase "uses . . . a firearm" is uttered does not preclude us from recognizing that there are other "uses" that qualify as well. In this case, it is both reasonable and normal to say that petitioner "used" his MAC-10 in his drug trafficking offense by trading it for cocaine; the dissent does not contend otherwise.

The dissent's example of how one might "use" a cane suffers from a similar flaw. To be sure, "use" as an adornment in a hallway is not the first "use" of a cane that comes to mind. But certainly it does not follow that the *only* "use" to which a cane might be put is assisting one's grandfather in walking.... [T]he use of a cane as an instrument of punishment was once so common that "to cane" has become a verb meaning "[t]o beat with a cane." In any event, the only question in this case is whether the phrase "uses . . . a firearm" in § 924(c)(1) is most reasonably read as *excluding* the use of a firearm in a gun-for-drugs trade. The fact that the phrase clearly *includes* using a firearm to shoot someone, as the dissent contends, does not answer it....

We are not persuaded that our construction of the phrase "uses . . . a firearm" will produce anomalous applications.... § 924(c)(1) requires not only that the defendant "use" the firearm, but also that he use it "during and in relation to" the drug trafficking crime. As a result, the defendant who "uses" a firearm to scratch his head, or for some other innocuous purpose, would avoid punishment for that conduct altogether: Although scratching one's head with a gun might constitute "use," that action cannot support punishment under § 924(c)(1) unless it facilitates or furthers the drug crime; that the firearm served to relieve an itch is not enough....

In any event, the "intended purpose" of a firearm is not that it be used in any offensive manner whatever, but rather that it be used in a particular fashion—by firing it. The dissent's contention therefore cannot be that the defendant must use the firearm "as a weapon," but rather that he must fire it or threaten to fire it, "as a gun." Under the dissent's approach, then, even the criminal who pistol-whips his victim has not used a firearm within the meaning of § 924(c)(1), for firearms are intended to be fired or brandished, not used as bludgeons. The universal

view of the courts of appeals, however, is directly to the contrary. No court of appeals ever has held that using a gun to pistol-whip a victim is anything but the "use" of a firearm; nor has any court ever held that trading a firearm for drugs falls short of being the "use" thereof. . . .

We therefore hold that a criminal who trades his firearm for drugs "uses" it during and in relation to a drug trafficking offense within the meaning of § 924(c)(1). Because the evidence in this case showed that petitioner "used" his MAC-10 machine gun and silencer in precisely such a manner, proposing to trade them for cocaine, petitioner properly was subjected to § 924(c)(1)'s 30-year mandatory minimum sentence. The judgment of the Court of Appeals, accordingly, is affirmed.

It is so ordered.

Justice Scalia, with whom Justice Stevens and Justice Souter join, dissenting: Section 924(c)(1) mandates a sentence enhancement for any defendant who "during and in relation to any crime of violence or drug trafficking crime . . . uses . . . a firearm." 18 U.S.C. § 924(c)(1). The Court begins its analysis by focusing upon the word "use" in this passage, and explaining that the dictionary definitions of that word are very broad. It is, however, a "fundamental principle of statutory construction (and, indeed, of language itself) that the meaning of a word cannot be determined in isolation, but must be drawn from the context in which it is used." That is particularly true of a word as elastic as "use," whose meanings range all the way from "to partake of" (as in "he uses tobacco") to "to be wont or accustomed" (as in "he used to smoke tobacco").

In the search for statutory meaning, we give nontechnical words and phrases their ordinary meaning. To use an instrumentality ordinarily means to use it for its intended purpose. When someone asks "Do you use a cane?," he is not inquiring whether you have your grandfather's silver-handled walking stick on display in the hall; he wants to know whether you *walk* with a cane. Similarly, to speak of "using a firearm" is to speak of using it for its distinctive purpose, *i.e.,* as a weapon. To be sure, "one can use a firearm in a number of ways," including as an article of exchange, just as one can "use" a cane as a hall decoration—but that is not the ordinary meaning of "using" the one or the other. The Court does not appear to grasp the distinction between how a word *can be* used and how it *ordinarily is* used. It would indeed, be "both reasonable and normal to say that petitioner 'used' his MAC-10 in his drug trafficking offense by trading it for cocaine." It would also be reasonable and normal to say that he "used" it to scratch his head. When one wishes to describe the action of employing the instrument of a firearm for such unusual purposes, "use" is assuredly a verb one could select. But that says nothing about whether the *ordinary* meaning of the phrase "uses a firearm" embraces such extraordinary employments. It is unquestionably *not* reasonable and normal, I think, to say simply "do not use firearms" when one means to prohibit selling or scratching with them. . . .

Given our rule that ordinary meaning governs, and given the ordinary meaning of "uses a firearm," it seems to me inconsequential that "the words 'as a weapon' appear nowhere in the statute"; they are reasonably implicit. Petitioner is not, I think, seeking to introduce an "additional requirement" into the text, but is simply construing the text according to its normal import.

The Court seeks to avoid this conclusion by referring to the next subsection of the statute, § 924(d), which does not employ the phrase "uses a firearm," but provides for the confiscation of firearms that are "used in" referenced offenses which include the crimes of transferring, selling, or transporting firearms in interstate commerce. The Court concludes from this that *whenever* the term appears in this statute, "use" of a firearm must include nonweapon use. I do not agree. We are dealing here not with a technical word or an "artfully defined" legal term, but with common words that are, as I have suggested, inordinately sensitive to context. Just as adding the direct object "a firearm" to the verb "use" *narrows* the meaning of that verb (it can no longer mean "partake of"), so also adding the modifier "in the offense of transferring, selling, or transporting firearms" to the phrase "use a firearm" *expands* the meaning of that phrase (it then includes, as it previously would not, nonweapon use). But neither the narrowing nor the expansion should logically be thought to apply to *all* appearance of the affected word or phrase. Just as every appearance of the word "use" in the statute need not be given the narrow meaning that word acquires in the phrase "use a firearm," so also

every appearance of the phrase "use a firearm" need not be given the expansive connotation that phrase acquires in the broader context "use a firearm in crimes such as unlawful sale of firearms." When, for example, the statute provides that its prohibition on certain transactions in firearms "shall not apply to the loan or rental of a firearm to any person for temporary use for lawful sporting purposes," I have no doubt that the "use" referred to is *only* use as a sporting *weapon*, and not the use of pawning the firearm to pay for a ski trip. Likewise when, in § 924(c)(1), the phrase "uses . . . a firearm" is not employed in a context that necessarily envisions the unusual "use" of a firearm as a commodity, the normally understood meaning of the phrase should prevail. . . .

Even if the reader does not consider the issue to be as clear as I do, he must at least acknowledge, I think, that it is eminently debatable—and that is enough, under the rule of lenity, to require finding for the petitioner here. . . .

For the foregoing reasons, I respectfully dissent.

REVIEW AND DISCUSSION QUESTIONS

1. What are the facts of the case?
2. According to O'Connor, what two things must the prosecution show? What is the "dispositive issue" in the case? If a word is not defined by statute, how should it be understood?
3. O'Connor argues that if Congress had intended a narrower construction ("use as a weapon" rather than simply "use"), then it could have said so. Since it did not, the Court cannot read the statute that way. Is this a reasonable principle of interpretation? What do you think Congress intended?
4. Scalia argues that the Court has failed to grasp the distinction between how a word *can be* used and how it *ordinarily is* used. Explain. How does O'Connor respond to this argument in her opinion?
5. As precedent, O'Connor points to cases of "pistol-whipping." What is her argument?
6. At the end of his dissent, Scalia appeals to the "rule of lenity." How persuasive do you find his argument?
7. Although Smith said he was willing to swap his gun for drugs, he never had a chance to do so. In your view, is he still guilty of the offense? Suppose the gun that Smith offered to swap had been left at home. Would this have mattered? What about a defendant who carries a concealed weapon to a drug deal, but never displays it or even mentions it? Has the firearm been "used" in connection with a drug trafficking crime?
8. His conviction on this charge brought Smith thirty years of imprisonment. Do you think that the legislature intended such a stiff penalty in cases likes this? Why do you think Congress passed this statute? Is it reasonable for a judge to bear in mind the stiffness of the sentence when determining whether Smith "used" a firearm?

Can a Murderer Inherit?

Riggs v. Palmer

Here the specific issue before the Court is whether a grandson who murdered his grandfather may nonetheless keep the inheritance provided in his grandfather's will. As with *Smith* and the Mann Act cases, however, there is deep disagreement between the judges over how the statute should be interpreted, as well as over the correct methods of statutory interpretation.

---

22 N.E. 188 (1889).

EARL, J.: On the 13th day of August, 1880, Francis B. Palmer made his last will and testament, in which he gave small legacies to his two daughters, Mrs. Riggs and Mrs. Preston, the plaintiffs in this action, and the remainder of his estate to his grandson, the defendant Elmer E. Palmer, subject to the support of Susan Palmer, his mother, with a gift over to the two daughters, subject to the support of Mrs. Palmer in case Elmer should survive him and die under age, unmarried, and without any issue. The testator, at the date of his will, owned a farm, and considerable personal property. . . . At the date of the will, and subsequently to the death of the testator, Elmer lived with him as a member of his family, and at his death was 16 years old. He knew of the provisions made in his favor in the will, and, that he might prevent his grandfather from revoking such provisions, which he had manifested some intention to do, and to obtain the speedy enjoyment and immediate possession of his property, he willfully murdered him by poisoning him. He now claims the property, and the sole question for our determination is, can he have it?

The defendants say that the testator is dead; that his will was made in due form, and has been admitted to probate, and that therefore it must have effect according to the letter of the law. It is quite true that statutes regulating the making, proof, and effect of wills and the devolution of property, if literally construed, and if their force and effect can in no way and under no circumstances be controlled or modified, give this property to the murderer. The purpose of those statutes was to enable testators to dispose of their estates to the objects of their bounty at death, and to carry into effect their final wishes legally expressed; and in considering and giving effect to them this purpose must be kept in view. It was the intention of the law-makers that the donees in a will should have the property given to them. But it never could have been their intention that a donee who murdered the testator to make the will operative should have any benefit under it. If such a case had been present to their minds, and it had been supposed necessary to make some provision of law to meet it, it cannot be doubted that they would have provided for it. It is a familiar canon of construction that a thing which is within the intention of the makers of a statute is as much within the statute as if it were within the letter; and a thing which is within the letter of the statute is not within the statute unless it be within the intention of the makers. The writers of laws do not always express their intention perfectly, but either exceed it or fall short of it, so that judges are to collect it from probable or rational conjectures only, and this is called "rational interpretation"; and Rutherford, in his Institutes (page 420), says: "Where we make use of rational interpretation, sometimes we restrain the meaning of the writer so as to take in less, and sometimes we extend or enlarge his meaning so as to take in more, than his words express." Such a construction ought to be put upon a statute as will best answer the intention which the makers had in view. . . .

> By an equitable construction a case not within the letter of a statute is sometimes holden to be within the meaning, because it is within the mischief for which a remedy is provided. The reason for such construction is that the law-makers could not set down every case in express terms. In order to form a right judgment whether a case be within the equity of a statute, it is a good way to suppose the law-maker present, and that you have asked him this question: Did you intend to comprehend this case? Then you must give yourself such answer as you imagine he, being an upright and reasonable man, would have given. If this be that he did mean to comprehend it, you may safely hold the case to be within the equity of the statute; for while you do no more than he would have done, you do not act contrary to the statute, but in conformity thereto. 9 Bac. Abr. 248.

In some cases the letter of a legislative act is restrained by an equitable construction; in others, it is enlarged; in others, the construction is contrary to the letter. . . . If the law-makers could, as to this case, be consulted, would they say that they intended by their general language that the property of a testator or of an ancestor should pass to one who had taken his life for the express purpose of getting his property? . . . [I]f an act of parliament gives a man power to try all causes that arise within his manor of Dale, yet, if a cause should arise in which he himself is party, the act is construed not to extend to that, because it is unreasonable that any man should determine his own quarrel. There was a statute in Bologna that whoever drew blood in

the streets should be severely punished, and yet it was held not to apply to the case of a barber who opened a vein in the street. It is commanded in the decalogue that no work shall be done upon the Sabbath, and yet giving the command a rational interpretation founded upon its design the Infallible Judge held that it did not prohibit works of necessity, charity, or benevolence on that day.

What could be more unreasonable than to suppose that it was the legislative intention in the general laws passed for the orderly, peaceable, and just devolution of property that they should have operation in favor of one who murdered his ancestor that he might speedily come into the possession of his estate? Such an intention is inconceivable. We need not, therefore, be much troubled by the general language contained in the laws. Besides, all laws, as well as all contracts, may be controlled in their operation and effect by general, fundamental maxims of the common law. No one shall be permitted to profit by his own fraud, or to take advantage of his own wrong, or to found any claim upon his own iniquity, or to acquire property by his own crime. These maxims are dictated by public policy, have their foundation in universal law administered in all civilized countries, and have nowhere been superseded by statutes. . . .

Here there was no certainty that this murderer would survive the testator, or that the testator would not change his will, and there was no certainty that he would get this property if nature was allowed to take its course. He therefore murdered the testator expressly to vest himself with an estate. Under such circumstances, what law, human or divine, will allow him to take the estate and enjoy the fruits of his crime? The will spoke and became operative at the death of the testator. He caused that death, and thus by his crime made it speak and have operation. Shall it speak and operate in his favor? If he had met the testator, and taken his property by force, he would have had no title to it. Shall he acquire title by murdering him? If he had gone to the testator's house, and by force compelled him, or by fraud or undue influence had induced him, to will him his property, the law would not allow him to hold it. But can he give effect and operation to a will by murder, and yet take the property? To answer these questions in the affirmative it seems to me would be a reproach to the jurisprudence of our state, and an offense against public policy. Under the civil law, evolved from the general principles of natural law and justice by many generations of juris-consults, philosophers, and statesmen, one cannot take property by inheritance or will from an ancestor or benefactor whom he has murdered. . . . Just before the murder he was not an heir, and it was not certain that he ever would be. He might have died before his grandfather, or might have been disinherited by him. He made himself an heir by the murder, and he seeks to take property as the fruit of his crime. What has before been said as to him as legatee applies to him with equal force as an heir. He cannot vest himself with title by crime. My view of this case does not inflict upon Elmer any greater or other punishment for his crime than the law specifies. It takes from him no property, but simply holds that he shall not acquire property by his crime, and thus be rewarded for its commission.

Our attention is called to *Owens v. Owens*, as a case quite like this. There a wife had been convicted of being an accessory before the fact to the murder of her husband, and it was held that she was nevertheless entitled to dower. I am unwilling to assent to the doctrine of that case. The statutes provide dower for a wife who has the misfortune to survive her husband, and thus lose his support and protection. It is clearly beyond their purpose to make provision for a wife who by her own crime makes herself a widow, and willfully and intentionally deprives herself of the support and protection of her husband. As she might have died before him, and thus never have been his widow, she cannot by her crime vest herself with an estate. The principle which lies at the bottom of the maxim *volenti non fit injuria* should be applied to such a case, and a widow should not, for the purpose of acquiring, as such, property rights, be permitted to allege a widowhood which she has wickedly and intentionally created.

. . . The judgment of the general term and that entered upon the report of the referee should therefore be reversed, and judgment should be entered as follows: That Elmer E. Palmer and the administrator be enjoined from using any of the personalty or real estate left by the testator for Elmer's benefit; that the devise and bequest in the will to Elmer be declared ineffective to pass the title to him; that by reason of

the crime of murder committed upon the grandfather he is deprived of any interest in the estate left by him; that the plaintiffs are the true owners of the real and personal estate left by the testator. . . .

GRAY, J., dissenting: [If] I believed that the decision of the question could be effected by considerations of an equitable nature, I should not hesitate to assent to views which commend themselves to the conscience. But the matter does not lie within the domain of conscience. We are bound by the rigid rules of law, which have been established by the legislature, and within the limits of which the determination of this question is confined. . . . Modern jurisprudence, in recognizing the right of the individual, under more or less restrictions, to dispose of his property after his death, subjects it to legislative control, both as to extent and as to mode of exercise. . . . To the statutory restraints which are imposed upon the disposition of one's property by will are added strict and systematic statutory rules for the execution, alteration, and revocation of the will, which must be, at least substantially, if not exactly, followed to insure validity and performance. . . . The capacity and the power of the individual to dispose of his property after death, and the mode by which that power can be exercised, are matters of which the legislature has assumed the entire control, and has undertaken to regulate with comprehensive particularity.

. . . I concede that rules of law which annul testamentary provisions made for the benefit of those who have become unworthy of them may be based on principles of equity and of natural justice. It is quite reasonable to suppose that a testator would revoke or alter his will, where his mind has been so angered and changed as to make him unwilling to have his will executed as it stood. But these principles only suggest sufficient reasons for the enactment of laws to meet such cases.

The statutes of this state have prescribed various ways in which a will may be altered or revoked; but the very provision defining the modes of alteration and revocation implies a prohibition of alteration or revocation in any other way. The words of the section of the statute are: "No will in writing, except in the cases hereinafter mentioned, nor any part thereof, shall be revoked or altered otherwise," etc. Where, therefore, none of the cases mentioned are met by the facts, and the revocation is not in the way described in the section, the will of the testator is unalterable. I think that a valid will must continue as a will always, unless revoked in the manner provided by the statutes. Mere intention to revoke a will does not have the effect of revocation. The intention to revoke is necessary to constitute the effective revocation of a will, but it must be demonstrated by one of the acts contemplated by the statute. As Woodworth, Jr., said in *Dan v. Brown,* 4 Oow. 490: "Revocation is an act of the mind, which must be demonstrated by some outward and visible sign of revocation." The same learned judge said in that case: "The rule is that if the testator lets the will stand until he dies, it is his will; if he does not suffer it to do so, it is not his will." . . . The finding of fact of the referee that presumably the testator would have altered his will had he known of his grandson's murderous intent cannot affect the question. We may concede it to the fullest extent; but still the cardinal objection is undisposed of—that the making and the revocation of a will are purely matters of statutory regulation, by which the court is bound in the determination of questions relating to these acts.

Two cases—in this state and in Kentucky—at an early day, seem to me to be much in point. *Gains v. Gains,* 2 A. K. Marsh. 190, was decided by the Kentucky court of appeals in 1820. It was there urged that the testator intended to have destroyed his will, and that he was forcibly prevented from doing so by the defendant in error or devisee; and it was insisted that the will, though not expressly, was thereby virtually, revoked. The court held, as the act concerning wills prescribed the manner in which a will might be revoked, that, as none of the acts evidencing revocation were done, the intention could not be substituted for the act. In that case the will was snatched away, and forcibly retained. In 1854, Surrogate Bradford, whose opinions are entitled to the highest consideration, decided the case of *Leaycraft v. Simmons,* 3 Bradf. Sur. 35. In that case the testator, a man of 89 years of age, desired to make a codicil to his will, in order to enlarge the provisions for his daughter. His son, having the custody of the instrument, and the one to be prejudiced by the change, refused to produce

the will at testator's request, for the purpose of alteration. The learned surrogate refers to the provisions of the civil law for such and other cases of unworthy conduct in the heir or legatee, and says: "Our statute has undertaken to prescribe the mode in which wills can be revoked [citing the statutory provision]. This is the law by which I am governed in passing upon questions touching the revocation of wills. The whole of this subject is now regulated by statute; and a mere intention to revoke, however well authenticated, or however defeated, is not sufficient." And he held that the will must be admitted to probate.

. . . The appellants' argument practically amounts to this: that, as the legatee has been guilty of a crime, by the commission of which he is placed in a position to sooner receive the benefits of the testamentary provision, his rights to the property should be forfeited, and he should be divested of his estate. To allow their argument to prevail would involve the diversion by the court of the testator's estate into the hands of persons whom, possibly enough, for all we know, the testator might not have chosen or desired as its recipients. Practically the court is asked to make another will for the testator. The laws do not warrant this judicial action, and mere presumption would not be strong enough to sustain it. But, more than this, to concede the appellants' views would involve the imposition of an additional punishment or penalty upon the respondent. What power or warrant have the courts to add to the respondent's penalties by depriving him of property? The law has punished him for his crime, and we may not say that it was an insufficient punishment. In the trial and punishment of the respondent the law has vindicated itself for the outrage which he committed, and further judicial utterance upon the subject of punishment or deprivation of rights is barred. We may not, in the language of the court in *People v. Thornton*, "enhance the pains, penalties, and forfeitures provided by law for the punishment of crime." The judgment should be affirmed, with costs.

REVIEW AND DISCUSSION QUESTIONS

1. Why does Judge Earl think Elmer should not inherit?
2. On what basis does Judge Gray conclude that the inheritance should be given to Elmer despite his crime?
3. Describe the different approaches the two judges take to statutory interpretation.
4. Judge Earl relies heavily on the notion of "principles" that although not explicit in any statute should nevertheless be given consideration when applying statutes to cases. Where does he think those principles originate? If challenged, how might he defend their soundness?
5. Does Judge Gray also rely on political principles? Explain.

CHAPTER 5

Precedents

Although the process of reasoning and argumentation used by lawyers and judges appears in some ways as a paradigm of analysis and argumentation, it differs strikingly from the form of reasoning that characterizes philosophy on the one hand or science on the other. For example, reliance on precedent (the decisions of previous judges), so distinctive of law, has no counterpart in other areas of intellectual inquiry. Precedents can take the form of either past decisions of judges who have interpreted a statute, or common law precedents that were created by judges in the absence of statutory law. The readings in this section discuss what is involved in following (or not following) precedent, as well as why it is important for the law to follow precedent.

Stare Decisis: The Use of Precedent

C. Gordon Post

Everyone knows that lawyers and judges rely on legal "precedents," but what precisely does that mean? In this essay, C. Gordon Post, professor of political science at Vassar College, discusses what precedents are, why following them is thought important, the ways judges eliminate old precedents and establish new ones, and the form of reasoning that is employed by judges when they appeal to a precedent. Along the way, he describes a wide variety of legal cases and opinions from both England and the United States that illuminate these important questions.

In the resolution of conflicts a court invokes and applies rules of law to proven facts. A question arises: Just where does a court find these rules?

There are two *chief* sources of law: *statutes* and *precedents.* The former, of course, come from the legislature, which consists of the elected representatives of the people. The latter come from the courts; precedents are the products of earlier decisions. To the latter, we should add the decisions of an increasing number of administrative bodies and the precedents established thereby. . . .

Most everybody knows what a statute is, but what is a precedent? In a general, nonlegal way, precedent plays an important role in our lives.

Often, we do things as our parents did them and cite their experience as precedent for what we do now; out of some continuing or repetitive situation there comes a rough rule of thumb. When a father is questioned as to why he spanked his son for some infraction of the household rules he might reply that as a boy in like circumstances he had been spanked as had his father before him. He might go on to explain that such treatment was an application of the rule of experience, "Spare the rod and spoil the child."

When a mother solicits the aid of her protesting children in doing the household chores, she might declare that children have always helped their parents, and in support of this experience she quotes the rule, "Many hands make light work." On the other hand, when the children clamor to help her make a cake for the church bazaar, she more than likely will decline the proffered aid and repeat the ancient adage, "Too many cooks spoil the broth."

A young man, intending a change from one job to another, is cautioned by his father, who tries to dissuade him with the rule, hoary with time, that "A rolling stone gathers no moss," or no money, as the case may be. The son, wanting a change of work, answers with another ancient rule, "Nothing ventured, nothing gained."

In all of these instances, there is the application of a rule of experience to a given situation. These are homely examples. Clubs, business organizations, boards of trustees, student groups—all have their rules, some written, some unwritten, which are often invoked as precedent for doing, or not doing, one thing or another. And a precedent here is defined by Webster as "something done or said that may serve as an example or rule to authorize or justify a subsequent act of the same or analogous kind."

JUDICIAL PRECEDENT

A judicial precedent is defined in the same dictionary as "a judicial decision, a form of proceeding, or course of action that serves as a rule for future determinations in similar or analogous cases."

The driver of a wagon loaded with buckskin goods stopped for the night at a certain inn. He was received as a guest and the innkeeper took charge of his property. During the night a fire broke out which resulted in the destruction of horses, wagon and goods. The owner of the property thus destroyed sued the innkeeper for damages.

Let us suppose that this was a case of first impression, that is, a situation which is before an American court for the first time. After hearing the evidence from both sides, the judge does not simply say, "I decide for the plaintiff," or "I decide for the defendant." He decides for one or the other and gives his reasons. He will speak as follows: "An innkeeper is responsible for the safe keeping of property committed to his custody by a guest. He is an insurer against loss, unless caused by the negligence or fraud of the guest, or by the act of God or the public enemy." The judge looks to the English common law and finds that the liability of innkeepers was expressed tersely in *Cross v. Andrews:* "The defendant, if he will keep an inn, ought, *at his peril,* to keep safely his guests' goods"; and at greater length by Coke in *Calye's Case:* "If one brings a bag or chest, etc., of evidences into the inn as obligations, deeds, or other specialities, and by default of the innkeeper they are taken away, the innkeeper shall answer for them."

The judge will go on to explain the reason for the rule. He will say that the rule has its origin in public policy.

> Every facility should be furnished for secure and convenient intercourse between different portions of the kingdom. The safeguards, of which the law gave assurance to the wayfarer, were akin to those which invested each English home with the legal security of a castle. The traveller was peculiarly exposed to depredation and fraud. He was compelled to repose confidence in a host, who was subject to constant temptation, and favored with peculiar opportunities, if he chose to betray his trust. The innkeeper was at liberty to fix his own compensation, and enforce summary payment. His lien, then as now, fastened upon the goods of his guest from the time they came to his custody. The care of the property was usually committed to servants, over whom the guest had no control, and who had no interest in its preservation, unless their employer was held responsible for its safety. In case of depredation by collusion, or of injury or destruction by neglect, the stranger would of necessity be at every possible disadvantage. He would be without the means either of proving guilt or detecting it. The witnesses to whom he must resort for information, if

not accessories to the injury, would ordinarily be in the interest of the innkeeper. The sufferer would be deprived, by the very wrong of which he complained, of the means of remaining to ascertain and enforce his rights, and redress would be well-nigh hopeless, but for the rule of law casting the loss on the party entrusted with the custody of the property, and paid for keeping it safely.[1]

STARE DECISIS

Let us suppose that a year or so later, another driver with a wagon load of hides spends the night at an inn. Again, the horses, the wagon, and the hides, are turned over to the innkeeper; and again, a fire occurs during the night and the property of the guest is burned up. The owner of the property then sues the innkeeper for damages. The situation here is exactly the same as in the earlier case.

The judge in the second case, according to the theory, will apply the rule or principle (which is the precedent) and decide in favor of the plaintiff. The precedent or authority of the first case is precise and fits the facts of the second case very nicely. This application by courts of rules announced in earlier decisions is spoken of as *stare decisis,* which means "let the decision stand." This has been, and is, a fundamental characteristic of the common law, although in American jurisdictions it is the practice upon occasion for a high court to overrule its own precedents.

Obviously, a legal system in which judges could decide cases any which way, manifesting prejudice, whimsy, ignorance and venality, each decision being an entity in itself unconnected with the theory, practices and precedents of the whole, would be a sorry system, or, one might say, no system at all, and a source of little comfort either to attorneys or litigants. Speaking of *stare decisis* many years ago, Judge Maxwell of the Supreme Court of Nebraska said: "In the application of the principles of the common law, where the precedents are unanimous in the support of a proposition, there is no safety but in a strict adherence to such precedents. If the court will not follow established rules, rights are sacrificed, and lawyers and litigants are left in doubt and uncertainty, while there is no certainty in regard to what, upon a given state of facts, the decision of the court will be."

One concludes, after a little thought, that *stare decisis* is "the instrument of *stability* in a legal system," that it "furnishes a legal system with *certainty* and *predictability*," and "clothes a legal system with *reliability*"; in addition, it "assures all persons of *equality and uniformity of treatment*" and judges with "an instrument of *convenience and expediency.*" In short, "*Stare decisis* preserves the judicial *experience* of the past."[2]

After a little more thought, however, one also sees that *stare decisis* is an instrument of conservatism, of immobility, of eyes-in-the-back-of-the-head, of stultification. The application of the same rule, decade after decade, long after changed conditions have robbed the rule of its validity, makes the rule a troublesome fiction.

American high courts do not hesitate to overrule their own precedents when social, economic, or political change demands a corresponding change in the law. Cardozo has said that "If we figure stability and progress as opposite poles, then at one pole we have the maxim of *stare decisis* and the method of decision by the tool of a deductive logic; at the other we have the method which subordinates origins to ends. The one emphasizes considerations of uniformity and symmetry, and follows fundamental conceptions to ultimate conclusions. The other gives freer play to considerations of equity and justice, and the value to society of the interests affected. The one searches for the analogy that is nearest in point of similarity, and adheres to it inflexibility. The other, in its choice of the analogy that shall govern, finds community of spirit more significant than resemblance in externals. 'Much of the administration of justice,' says Pound, 'is a compromise between the tendency to treat each case as one of a generalized type of case, and the tendency to treat each case as unique.' Each method," concludes Cardozo, "has its value, and for each in the changes of litigation there will come the hour for use. A wise eclecticism employs them both."[3]

PRECEDENT AND FACTS

In the two cases outlined above, it has been assumed that the facts were identical, but only rarely are fact situations exactly alike. And from a different fact situation there may arise a different

precedent or rule of law. Professor Rodell, taking as an example a collision of two automobiles, points out the difficulties:

> Suppose a man driving a 1939 Cadillac along the Lincoln Highway toward Chicago runs into a Model T Ford, driven by a farmer who has just turned onto the Highway from a dirt road, and demolishes the Ford but does not hurt the farmer. The farmer sues, and a local judge, on the basis of various principles of Law which are said to "control" the case, awards him $100. A week later, another man driving a 1939 Cadillac along the Lincoln Highway toward Chicago runs into a Model T Ford driven by another farmer who has just turned onto the Highway from the same dirt road, and demolishes the Ford but does not hurt the farmer. This farmer also sues. The facts, as stated, seem to make this case quite similar to the previous case. Will it then fall into the same group of fact situations? Will it be "controlled" by the same principles of Law? Will the second farmer get $100?
>
> That all depends. For of course [as in the innkeeper's case] there will be other facts in both cases. Some may still be similar. Others, inevitably, will be different. And the possibilities of variation are literally endless.
>
> Maybe the first Cadillac was doing sixty miles an hour and the second one thirty. Or maybe one was doing forty-five and the other one forty. Or maybe both were doing forty-five but it was raining one week and clear the next. Maybe one farmer blew his horn and the other didn't. Maybe one farmer stopped at the crossing and the other didn't. Maybe one farmer had a driver's license and the other didn't. Maybe one farmer was young and the other was old and wore glasses. Maybe they both wore glasses but one was nearsighted and the other farsighted.
>
> Maybe one Cadillac carried an out-of-state license plate and the other a local license plate. Maybe one of the Cadillac drivers was a bond salesman and the other a doctor. Maybe one was insured and the other wasn't. Maybe one had a girl in the seat beside him and the other didn't. Maybe they both had girls beside them but one was talking to his girl and the other wasn't.
>
> Maybe one Cadillac hit its Ford in the rear left wheel and the other in the front left wheel. Maybe a boy on a bicycle was riding along the Highway at one time but not the other. Maybe a tree at the intersection had come into leaf since the first accident. Maybe a go-slow sign had blown over.[4]

In the second case of the innkeeper and the guest a difference in the fact situation, no matter how similar it may be to that of the first case, may lead a court to *distinguish* the one case from the other and to apply therefore a different precedent or rule.

James P. Tallon was employed as a guard on one of the trains of the Interborough Rapid Transit Company. He lived at 146th Street in Manhattan and it was his duty to report for duty each morning at the Transit Company's station at 177th Street and Third Avenue. To get to and from work, he was provided with a pass by the company which allowed him free passage. One morning on his way to work and in the uniform of a guard, a collision of trains occurred which resulted in Tallon's death.

Tallon's wife brought suit against the company for damages, alleging negligence on the part of the company, and claiming that her husband was a passenger on one of the defendant's trains and not an employee within the meaning of the Workmen's Compensation Law. [If he were traveling as an employee, then he could not sue but instead could only collect workmens' compensation.]

The defendant insisted in the trial that, "Tallon was injured in an accident arising out of and in the course of his employment and that this action could not be maintained." The trial judge ruled as a matter of law that Tallon had been a passenger and "the negligence being admitted, left to the jury solely the question of damages."

The Appellate Division reversed the judgment of the trial court and dismissed the complaint on the ground that the only relief for Mrs. Tallon was under the Workmen's Compensation Law, since Tallon, at the time of his death, was within the terms of that law in that "he was injured in an accident arising out of and in the course of his employment."

The Court of Appeals, speaking through Judge Crane, reversed the Appellate Division and affirmed the judgment of the trial court. Judge McLaughlin, dissenting, and supporting the judgment of the intermediate court, was of the opinion that "at the time of the collision Tallon occupied the status of an employee and, therefore, relief should have been sought under the Workmen's Compensation Law." At the time of the accident he was riding on a pass which entitled him to free transportation to and from his work. Such transportation was an incident of the employment. It was a part of the contract of

employment and enforceable by him as such. The facts bring the case directly within the principle laid down in *Matter of Littler v. Fuller Co.*

Littler was a bricklayer. He worked for the Fuller Co., which was constructing a residence in Great Neck, some two miles from the railroad station. The workmen who came out to Great Neck by train had refused to continue to work unless the employer would furnish them with free transportation from the railroad station and back again at the end of a day's work. The employer hired a truck to transport the men. Littler was injured when the truck, making its way back to the station, turned over in a ditch.

Littler sued the Fuller Company, arguing that his injury did not arise out of or in the course of employment. The Court of Appeals thought otherwise. "The vehicle," said Judge Pound, "was provided by the employer for the specific purpose of carrying the workmen to and from the place of the employment and in order to secure their services. The place of injury was brought within the scope of the employment because Littler, when he was injured, was 'on his way . . . from his duty within the precincts of the company.' . . . The day's work began when he entered the automobile truck in the morning and ended when he left it in the evening. The rule is well established that in such cases compensation should be awarded." Judge Pound goes on to say that the "case would be different if at the time of the accident" Littler "had been on the railroad train on his way to or from Great Neck."

Judge Crane, speaking for the majority in the Tallon case, *distinguished* the Littler case from the matter at hand. He said,

> The cardinal underlying fact is that Tallon's employment did not actually begin until he reported for work at One Hundred and Seventy-seventh Street and Third avenue. He had to get there, and get there on time, and to facilitate his arriving on time the defendant gave him the right to ride in its passenger trains free of charge, but I cannot see how this in any way changes the reality, the existing fact, that the employment commenced at One Hundred and Seventy-seventh Street and Third avenue at the time of reporting. It would cause no such change, had Tallon paid his fare on defendant's train or had ridden in another conveyance, the defendant paying his fare. The pass alone, even though it be a part of his compensation, cannot create a fictitious relationship.

But there is still the Littler case to be considered, which Judge McLaughlin in his dissent thought applicable in this matter. Judge Crane, therefore, *distinguished* the Littler case as follows: "Now this case differs materially from those cases where the employer in order to get his employees to and from their work, provide conveyances exclusively for their use which in no sense are public conveyances and in which the employees undertaken to ride as part of their contract of employment in going to and from their work. Such a case was *Matter of Littler v. Fuller Co.*"

Thus, Littler had to accept an award under Workmen's Compensation while Mrs. Tallon would receive damages from the wealthy construction company. Between the majority view and the dissenting view of the facts there was a decided difference; each view elicited a different precedent and each produced a different result.

IMPRECISE PRECEDENTS

The Tallon matter might be called a run-of-the-mill case, and it is an example of the kind of thing courts do every day. They "distinguish" or they "follow" precedents. But what happens when the precedents do not fit the facts or two precedents might well and equally, but not exactly, cover a situation?

It was customary some years ago for young boys to swim in the United States Ship Canal, a narrow body of water separating Manhattan and the Bronx. They would cross from the Manhattan side and climb the bulkhead belonging to the New York Central Railroad and to which was attached, by whom or when no one knew, a rough plank which served as a diving board. This plank extended several feet beyond the line of the railroad's property. As sixteen-year-old Harvey Hynes stood poised on the board ready to dive, a high tension wire belonging to the railroad fell upon him and swept him to his death below. Hynes' mother brought an action for damages against the company.

The complaint brought by Mrs. Hynes alleged that the defendant's neglect in improperly erecting, constructing, and maintaining its poles and appurtenances and the wires attached thereto was responsible for the death of her son.

In the court of first instance, a jury awarded the plaintiff, Mrs. Hynes, eight thousand dollars. The defendant, the railroad, at once appealed to the Supreme Court of New York to set aside the verdict of the jury and to grant a new trial, and the Supreme Court so ordered. Mrs. Hynes then appealed from this order to the Appellate Division of the Supreme Court.

Before the Appellate Division, the defendant argued that Hynes, in climbing the bulkhead and standing in readiness to plunge from the diving board into the water, was a trespasser. True, the railroad was duty bound "to use reasonable care that bathers swimming or standing in the water should not be electrocuted by wires falling from its right of way," but that there was no duty to a boy on the springboard unless any injury he might receive was "the product of mere wilfulness or wantonness."

The court accepted his argument. The majority held that one who trespasses on a riparian owner's land and goes from there onto a springboard extending out over public waters is still a trespasser even though he is on that part of the board beyond the technical boundary of the riparian owner.

Two of the Appellate Division judges did not agree with the majority and one of them wrote a dissenting opinion. His argument was that Hynes was not a trespasser in the usual sense. Ordinarily, in like cases, the accident would not have happened but for the trespass. In this case, however, the falling wires would have been as fatal to the boy swimming in the river at that point as standing upon a plank attached to the defendant's premises. "If a wagon of the defendant," he said, "has been left standing in the highway" and young Hynes "had climbed into that wagon and while there had been killed by the defendant's wires falling into the highway, there would have been no question as to the plaintiff's right to recover. The situation," he continued, "is exactly analogous to the situation involved in this action. The boy's death was not caused by a defect in the premises upon which he trespassed, but by the negligence of the defendant which permitted its wires to fall into the navigable waters of the river."

Finally, Mrs. Hynes appealed to the Court of Appeals, the highest court of the State of New York. Here, the decision of the Appellate Division was reversed.

In the lower courts, the judges emphasized the concept of the trespasser. Hynes was a trespasser pure and simple and therefore his mother could not recover. The Court of Appeals, speaking through Judge Cardozo, gave equal emphasis to another concept, namely, the concept of the traveller on the highway. This concept had its beginning at a time when roads were not lighted and when it was quite possible on a moonless night for the traveller to stray from the highway. If, for example, a landowner dug a hole four or five feet from the highway and the traveller strayed a little off the road at night, fell into the hole, and was injured, the landowner would be liable in damages even though the traveller was an inadvertent trespasser. To this day, one sees with the coming of darkness red lanterns and planks surrounding any kind of excavation near the highway. If the wayfarer fell into a hole three hundred yards from the road, it might be surmised that the luckless one was there for no good reason or that he was in such a state as not to know the difference between a highway and a melon patch. The traveller concept would hardly apply.

"Landowners," said Judge Cardozo, "*are* bound to regulate their conduct in contemplation of the presence of travellers upon the adjacent public ways." Hynes, until the moment of his death, was in the enjoyment of the public waters and when he climbed the bulkhead and stood on the board poised to dive, he had not abandoned his "rights as a bather." The court chose to look upon the use of the springboard as "a mere byplay, an incident, subordinate and ancillary to the execution of his primary purpose, the enjoyment of the highway. The byplay, the incident, was not the *cause* of the disaster. Hynes would have gone to his death if he had been below the springboard or beside it."

The railroad would have been liable if the wires had fallen upon a bather in the water, or upon a fisherman in a rowboat, or upon the pilot of an airplane skimming a few feet above the public waters. But because the board was attached to land belonging to the railroad, it was deemed in the lower court to be an extension of the company's property and the use of it by the boy, a trespass. The Court of Appeals could not accept this view.

"Two boys walking in the country or swimming in a river," said Judge Cardozo, "stop to rest for a moment along the side of the road or the margin of the stream. One of them throws himself beneath the overhanging branches of a tree. The other perches himself on a bough a foot or so above the ground. . . . Both are killed by falling wires. The defendant would have us say that there is a remedy for the representatives of one, and none for the representatives of the other."

Cardozo states the problem succinctly: in one sense, Hynes standing on the springboard was a trespasser; in another sense, he was still on the public waters in the enjoyment of public rights. "The law," said Cardozo, "must say whether it will subject him to the rule of the one field or of the other. . . . "

Considerations "of analogy, of convenience, of policy, and of justice," led the Court of Appeals to hold the railroad "in the field of liability and duty."

Commenting later on the analogy of the trespasser and the analogy of the traveller on the highway, Cardozo said that "as a mere bit of dialectics, these analogies would bring a judge to an impasse. No process of merely logical deduction could determine the choice between them. Neither analogy is precise, though each is apposite. There had arisen a new situation which could not force itself without mutilation into any of the existing moulds. When we find a situation of this kind," he continued, "the choice that will approve itself to this judge or to that, will be determined largely by his conception of the end of the law, the function of legal liability; and this question of ends and functions is a question of philosophy."[5] . . .

THE CREATION OF NEW PRECEDENTS

Stare decisis is a characteristic of the common law; but it is true that in American jurisdictions high courts will, when conditions change or necessity demands, overrule their own precedents.

Robert C. Woods was *en ventre sa mere* in the ninth month of his mother's pregnancy when she fell down a stairway in a multiple dwelling house, causing injuries to herself and to the child who was born thirteen days after the accident. The child came into the world permanently maimed and disabled. The mother, in behalf of her infant son, sued the owner of the multiple dwelling, the complaint alleging negligence.

The suit was brought in the Supreme Court for Bronx County. The defendant moved for a dismissal of the complaint on the ground that it failed to state a cause of action. The court sustained the motion and dismissed the complaint. From this order, the plaintiff appealed to the Appellate Division of the Supreme Court which affirmed (four to one) the order of the lower court.

The four judges who affirmed the order wrote no opinion, except for Judge Shientag, who concurred but would have allowed recovery "provided the causal relation between the negligence and the damage to the child be established by competent medical evidence"; he concurred with the majority because thirty years earlier the Court of Appeals had squarely decided against the plaintiff.

This was in the case of *Drobner v. Peters*, decided in 1921. The defendant negligently permitted a coal hole in the sidewalk in front of his premises to remain uncovered and the plaintiff's mother fell into it. The child was born eleven days later. He had sustained injuries as a result of the accident. The question before the Court of Appeals was: "Does the complaint herein state facts sufficient to constitute a cause of action?" It was contended by the defendant that at the time of the injury the child "was not a person but was a part of the body of his mother and that, as the injury was to the mother, he has no cause of action." The Court of Appeals answered the question in the negative.

To return to *Woods v. Lancet*, Judge Heffernan (in the Appellate Division of the Supreme Court) dissented. He referred to *Drobner v. Peters* and admitted that no case could be found in New York wherein a plaintiff had been awarded damages for injuries inflicted upon his person while *in utero.* He said, however, that "an adjudicated case is not indispensable to establish a right to recover under the rules of the common law. Precedent is merely a guide; its absence never a bar. The fact that there is no precedent is not conclusive. The law would be an absurd science were it founded on precedents only. The absence of precedent should give no immunity to one who by his wrongful act

has invaded the right of an individual. No right is more inherent, more sacrosanct, than that of an individual in his possession and enjoyment of his life, his limbs and his body. The law is not static and inert, but is sufficiently elastic to meet changing conditions. It is presumed to keep pace with present day concepts. To deny the infant relief in this case is not only a harsh result, but its effect is to do reverence to an outmoded, timeworn fiction not founded on fact and within common knowledge untrue and unjustified."

Judge Heffernan referred to the fact that courts in other jurisdictions had held "that an unborn viable child is capable of independent existence and" hence "should be regarded as a separate entity." . . .

The Court of Appeals in *Woods v. Lancet* reversed both lower courts. Speaking through Judge Desmond, a majority held that recovery should be allowed. In his opinion, Judge Desmond reviewed the history of prenatal injury, going back no further than *Dietrich v. Northampton*, decided in 1884, and "apparently, the first American case." In an opinion by Justice Holmes, the Massachusetts Supreme Court denied recovery on the grounds that "the unborn child was a part of the mother at the time of the injury," and that any damage to it which was not too remote to be recovered for at all was recoverable by her. Judge Desmond cited other cases where courts had refused to allow a suit to be brought on behalf of a child who sustained an injury prior to birth.

In *Drobner v. Peters*, the Court of Appeals, finding no precedent for allowing the suit, had adopted the general theory of *Dietrich v. Northampton.* Judge Desmond went on to say that since 1921, "numerous and impressive affirmative precedents have been developed." The highest courts in California, Ohio, Minnesota, Maryland, and Georgia, had all upheld the right of an infant to bring an action for prepartum injuries tortiously inflicted. "What, then, stands in the way of reversal here?" asked Judge Desmond. "Surely", he said, "as an original proposition, we would, today, be hard put to it to find a sound reason for the old rule. Following *Drobner v. Peters* . . . would call for an affirmance but the chief basis for that holding (lack of precedent) no longer exists."

The argument as to the supposed difficulty of proving or disproving that certain injuries befell the unborn child is a valid one but does not present an insuperable obstacle since "Every day in all our trial courts (and before administrative tribunals, particularly the Workmen's Compensation Board), such issues are disposed of, and it is an inadmissible concept that uncertainty of proof can ever destroy a legal right."

Thus *Drobner v. Peters* was overruled and recovery was allowed. Anyone bringing suit for injuries before birth must still *prove* that the injuries were the result of the accident and tortiously inflicted.

CERTAINTY AND THE LAW

The overruling of ancient precedents is a cause of dismay to many lawyers; and those who are not lawyers tend to look askance at a much-vaunted legal system which allows such a degree of uncertainty. "The law ought to be certain," is a remark often heard. "How can I know my rights if the law itself is uncertain?"

The desire for certainty is all-important. We like to feel certain of an economic system which gives us security and a political system which safeguards that security. We like the comfort of the absolute.

Yet individuals too readily forget that their forefathers once subscribed with absolute certainty to astrology, the existence of witches and bogey-men, and the notion that the earth was a flat body; individuals also forget that the present is very much a part of the historical process and look upon society, though turbulent in its development, as static.

If history teaches us anything, it is that society is not static. This lesson, simple as it is, and implied on every page of the most elementary textbook, is seemingly lost, as it always has been lost, in the illusion that as things are so shall they ever be; we wait until new social conditions, long unrecognized or ignored, pull the rug under our feet and force us to a revision of our theories, or our precedents. "The logical method and form," declared Mr. Justice Holmes, "flatter that longing for certainty and for repose which is in every human mind. But certainty generally is an illusion, and repose is not the destiny of man." . . .

REASONING BY ANALOGY

Donald MacPherson had bought a Buick automobile from a dealer in Schenectady, New York. While in the car with two friends and travelling at the rate of about eight miles an hour, the spokes of the left rear wheel collapsed and Mr. MacPherson was thrown into a ditch and injured. He brought suit, not against the dealer with whom he had entered into a contract for the purchase of the car, but against the Buick Motor Company, with whom he had no contractual relationship whatever. The general rule was well expressed by Chief Justice Cooley of Michigan many years ago: "The general rule is that a contractor, manufacturer, vendor or furnisher of an article is not liable to third parties who have no contractual relations with him for negligence in the construction, manufacture or sale of such article." And the exceptions to the general rule recognized in New York are cases in which the article sold "was of such a character that danger to life or limb was involved in the ordinary use thereof; in other words, where the article sold was inherently dangerous."

The trial judge instructed the jury that an automobile was not an inherently dangerous vehicle but that it may become so if it is equipped with a defective wheel; and that "if the motor car in question, when it was put upon the market was in itself inherently dangerous by reason of its being equipped with a weak wheel, the defendant was chargeable with a knowledge of the defect so far as it might be discovered by a reasonable inspection and the application of reasonable tests." The jury returned a verdict for the plaintiff for $5,025.

The Buick Company then appealed to the Appellate Division of the Supreme Court. This court affirmed the judgment of the lower court. . . .

Judge Kellogg stressed the fact that a defective wheel is liable to cause serious damage not only to the people using the car but to others as well; an automobile, he said, "equipped with weak wheels" is an "imminently dangerous machine."

Speaking *obiter,* Judge Kellogg said that

> In the old days a farmer who desired to have wheels made for an ox-cart would be apt to inspect the timber before it was painted, before the wheel was ironed and the defects covered up, in order that he might know what he was buying. He would realize that the oxen, in case of an accident or fright, as he would say, "might go pretty fast," and that if a wheel broke serious damage might occur to him or to others. . . . An ordinary man, in buying a pitchfork, a golf club, an axe-helve, or an oar for a boat, will look at the timber, "heft it," and otherwise endeavor to ascertain whether it is made of suitable material.

Since the ultimate purchaser is not qualified to make such an examination, or even to examine the quality of the wood, it is incumbent upon the manufacturer to do it for him.

The court held that under the circumstances, the Buick Company "owed a duty to all purchasers of its automobiles to make a reasonable inspection and test to ascertain whether the wheels purchased and put in use by it were reasonably fit for the purposes for which it used them, and if it fails to exercise care in that respect it is responsible for any defect which would have been discovered by such reasonable inspection and test."

The Buick Company appealed to the New York Court of Appeals. The sole question to be determined was "whether the defendant owed a duty of care and vigilance to anyone but the immediate purchaser," that is, the retail dealer. . . .

The general rule, you will remember, was well expressed by Judge Cooley. An exception to the general rule is the doctrine, if such it may be called, of the "inherently dangerous instrument." If the article is "inherently" or "imminently" dangerous, then the manufacturer is liable to third parties; if it is not within the category of "inherently dangerous," then the manufacturer is not liable to third parties. The Buick Company struggled to the bitter end to persuade the court not to include an automobile in this category, while the respondent sought to persuade the court to keep the car in the category in line with the lower court opinion.

The Court of Appeals, speaking through Judge Cardozo, affirmed the judgment of the Appellate Division. First he reviewed the precedents. That a manufacturer may be liable to a third party was established in New York in the case of *Thomas v. Winchester,* decided in 1853. A druggist had purchased a bottle of what was supposedly extract of dandelion from Winchester, who operated a pharmaceutical business in New York City. The bottle,

through the negligence of one of Winchester's employees, had been incorrectly labeled extract of dandelion although the bottle did in fact contain belladonna, a poison. A Mr. Thomas bought a small quantity of this liquid from the druggist for his wife, who, after taking a few spoonfuls, became violently sick. Mrs. Thomas sued Winchester, not the druggist, and recovered damages. "The defendant's negligence," said the Court, "put human life in imminent danger." With regard to this early case, Cardozo said, "A poison falsely labeled is likely to injure anyone who gets it. Because the danger is to be foreseen, there is a duty to avoid the injury."

In 1882, in *Devlin v. Smith,* the Court of Appeals had under consideration a case in which the defendant, a contractor, contracted to build a scaffold for a painter. The painter's workmen were injured when the scaffold collapsed, and the contractor was held liable although there was no privity of contract between him and the workmen. The defendant, said Cardozo, "knew that the scaffold, if improperly constructed, was a most dangerous trap. He knew that it was to be used by the workmen. . . . Building it for their use, he owed them a duty, irrespective of his contract with their master, to build it with care."

Cardozo then considers another case in which the rule of *Thomas v. Winchester* was followed: *Statler v. Ray Manufacturing Company.* Here the manufacturer of a coffee urn was held liable when one of his urns, installed in a restaurant, exploded and injured the plaintiff. The court declared that the urn "was of such a character inherently that, when applied to the purposes for which it was designed, it was liable to become a source of great danger to many people if not carefully and properly constructed."

To the argument that things imminently dangerous to life were poisons, explosives, and deadly weapons, and that (reasoning by analogy) the automobile was not like any of these, Cardozo said: "Whatever the rule in *Thomas v. Winchester* may once have been, it has no longer that restricted meaning. A scaffold . . . is not inherently a destructive instrument. It becomes destructive only if imperfectly constructed. A large coffee urn . . . may have within itself, if negligently made, the potency of danger, yet no one thinks of it as an implement whose normal function is destruction." After a review of New York court decisions, Judge Cardozo declared that the rule of *Thomas v. Winchester* was not limited to "poisons, explosives, and things of like nature. . . . " He goes on to state the rule: "If the nature of a thing is such that it is reasonably certain to place life and limb in peril when negligently made, it is then a thing of danger. Its nature gives warning of the consequences to be expected. If to the element of danger there is added knowledge that the thing will be used by persons other than the purchaser, and used without new tests, then, irrespective of contract, the manufacturer of this thing of danger is under a duty to make it carefully. That is as far as we are required to go for the decision of this case."

Cardozo then goes on to clarify the rule. "There must be knowledge of a danger," he said, "not merely possible, but probable. It is *possible* to use almost anything in a way that will make it dangerous if defective. That is not enough to charge the manufacturer with a duty independent of his contract." . . .

The MacPherson and Smith cases, as compared with *Thomas v. Winchester,* reveal an enlargement of a rule of society as manifested by the courts. Certainly, in this area the law was brought more "in line with social considerations." Unlike a hundred years ago, the enlarged ruling recognized the new position of the buyer vis-à-vis the manufacturer. The meat, the canned or frozen fruits and vegetables, the bottled waters—these come from distant places and merely pass through the hands of the retailer to the consumer. If a young man buys a bottle of ginger beer from a retailer for his girl and she drinks the concoction only to find the remains of a decomposed snail in the bottom of the bottle and becomes deathly sick, are we, in this day and age, to say that the poisoned consumer has no remedy against the negligent manufacturer?

As Mr. Holmes said, with some tenderness for a felicitous statement, the life of the law has not been logic, it has been experience. Perhaps this is not altogether true, but the cases do reflect the changing attitudes of courts to the changing facts of social life. And, as Professor Levi says, to contrast "logic and the actual legal method [reasoning by example] is" to do "a disservice to both. Legal reasoning has a logic of its own. Its structure fits it to give meaning to ambiguity and to test constantly whether the society has come to see new differences or similarities.

Social theories and other changes in society will be relevant when the ambiguity has to be resolved for a particular case."[6]

In the absence of statute, courts as a general rule will follow precedent or the rules announced by judges in earlier cases. As often as not, new facts do not adapt themselves precisely to existing precedent and, therefore, courts must make distinctions, analogies must be drawn, and in some cases precedents are overruled and new ones created. Where the legislature by statute moves in on territory occupied by precedent, the statute prevails. . . . But statutes are not always crystal-clear and admit of various interpretations which may even allow the continuance of a common law rule it was presumably the intention of the legislature to abolish.

NOTES

1. *Hulett v. Swift* (1865) 33 N.Y. 572–573.
2. Robert A. Sprecher, "The Development of the Doctrine of *Stare Decisis* and the Extent to which it should be Applied," in 31 *American Bar Association Journal* (Oct., 1945), p. 501.
3. Benjamin N. Cardozo, *The Parodoxes of Legal Science* (New York: Columbia University Press, 1928), p. 8.
4. Fred Rodell, *Woe Unto You, Lawyers!* (New York: Pageant Press, 1957), pp. 116–118.
5. Benjamin N. Cardozo, *The Growth of the Law* (New Haven: Yale University Press, 1927), pp. 100–101.
6. Edward H. Levi, *An Introduction to Legal Reasoning* (Chicago: The University of Chicago Press, 1949), pp. 73–74. In the foregoing section, the writer is greatly indebted to Professor Levi's stimulating and enlightening study.

REVIEW AND DISCUSSION QUESTIONS

1. Explain the various reasons for thinking *stare decisis* is important.
2. As Post explains, precedents can be (1) distinguished, (2) overturned, or (3) revised (limited or expanded). Give examples of each of these from his essay.
3. Given the importance of following precedents, why may it be important to revise or overturn them?

Constrained by Precedent

Larry Alexander

According to the doctrine of precedent or *stare decisis*, courts should follow prior legal decisions. Although precedent does not override every other legal consideration, precedent does constrain a court's decision. In this essay, Larry Alexander presents three different models for explaining or justifying precedential constraint—the natural model, the rule model, and the pure result model—examining the strength and weaknesses of each and comparing their merits.

I. THE FIRST MODEL: GIVING PRECEDENT ITS NATURAL WEIGHT

A. The Natural Model of Precedent Described

The first model of precedent following is what I call the natural model. Under this model the court in deciding a case gives prior judicial decisions the weight that those decisions carry independently of any formal requirement that precedent be followed. In other words, the constrained court looks to the "precedential" effects of the prior decision and assigns them a moral weight, which it then factors into its overall decisionmaking calculus.

To understand the natural model better, let us look at an instance of nonlegal decisionmaking where the model clearly does operate. Suppose that when my daughter reaches the age of thirteen,

she requests permission from me to attend a rock concert. I weigh the possible risks involved in her going and the potential benefits to her and to our relationship, and I decide to grant her permission to go. When my son reaches the age of thirteen and also seeks permission to attend a rock concert, he predictably will cite my previous decision granting permission to his sister as a reason for a decision in his favor, a reason in addition to the risks and benefits I would normally consider. The natural model of precedent requires that I give the previous decision its true weight as an independent reason.

The natural model would have the same features in the legal system. In the precedent case, which I am assuming is not governed by any statute or constitutional provision that speaks to its substantive merits, and which I shall also assume is not itself governed by an applicable precedent case, the court—which I shall hereafter call the "precedent court"—will assess the reasons why the dispute should be resolved in one party's favor or the other's. Those reasons might primarily be the parties' relative moral desert or some other deontological consideration—a sort of "What if these parties were alone on a desert island?" inquiry. Alternatively, those reasons might primarily be the social consequences of the alternative decisions. Although it is interesting to reflect upon how a court should resolve a dispute in the absence of controlling statutory rules or precedent, we do not need to assume anything beyond the fact that the precedent court does in fact resolve the dispute based on reasons—deontological, consequentialist, or some mixture—that it believes correct political/moral theory approves.

The important point here is that once the precedent court decides the case, the existence of that decision can be invoked as an independent reason for subsequent courts to decide "similar" disputes in the same way. This is true even in the absence of a formal practice of precedent following. Similarly, my decision to allow my daughter to attend a rock concert at age thirteen can (and will) be invoked by my son as a reason for permitting him to attend a rock concert at the same age. The earlier decision is invoked as a reason for a similar decision, a reason that did not exist in the precedent case and that is to be added to the other reasons for such a decision. More precisely, the precedent decision gives rise to reasons for a similar decision, such as equality and reliance, that did not exist in the precedent case.

Moreover, even if the subsequent court—which I shall hereafter call the "constrained court"—believes the precedent case was decided incorrectly because the balance of reasons favored the losing party, the constrained court might believe that the new reasons generated by the precedent decision tip the balance in favor of the party who would have lost had the constrained court decided the precedent case. Similarly, I might decide that although I should not have let my daughter attend the rock concert, and should not let my son attend were this my first decision on the point, the decision in my daughter's case new tips the balance of reasons in favor of permitting my son to go.

B. The Rationale Behind the Natural Model of Precedent

I suspect that we all recognize the natural model, if not in the legal system, at least in other areas of our lives. But how can a past decision that we now believe was incorrect convert a present decision that would otherwise be incorrect into a correct decision? The arguments on behalf of following incorrect decisions on the natural model of precedent boil down to two reasons: furthering equality of treatment and respecting justified expectations on which people have relied.

It is easy to see why one might argue that equality supports the constrained court's decision that, but for the precedent court's decision, would be incorrect. If I tell my son that I will not permit him to attend the rock concert and that I should not have permitted his sisters to go, he predictably will claim that I have not treated him equally with respect to her. By saying so, he will be making more than the obvious descriptive point that my treatment of his sister and him with respect to attending rock concerts at age thirteen is dissimilar. Rather, he will be making the normative claim that such dissimilarity in treatment is wrong in the absence of countervailing considerations of substantial moral weight. In other words, he will be asserting that the value of equality is a sufficiently weighty reason in support of letting him go to the concert that it tips the balance of reasons in his favor.

The role of the justified reliance argument in support of a decision in favor of the present analog to the party who won in the precedent case is also fairly easy to understand. Even if there is no

formal requirement that courts follow precedent, people frequently will look to judicial decisions and the expressed reasons on which they are based in order to predict how their contemplated courses of action will be treated by the courts should a legal dispute arise. Similarly, my son may undertake action based upon his reasonable expectation that because I allowed his sister to attend a rock concert at age thirteen, I will allow him to do the same. These expectations become reasons that may justify decisions that would otherwise (in the absence of the expectations) be regarded as erroneous. . . .

II. THE SECOND MODEL: THE RULE MODEL OF PRECEDENT

A. The Rule Model of Precedent Described

The second model of precedential constraint is the rule model of precedent. Under the rule model, the precedent court has authority not only to decide the case before it but also to promulgate a general rule binding on courts of subordinate and equal rank. The rule will operate like a statute and will, like a statute, have a canonical formulation.

Versions of the rule model can very along two dimensions. First, they can vary according to the strength by which the precedent court's rule binds the constrained court. On some versions, the constraint may be absolute, like that of a statute or constitutional rule on a court. On those versions, the constrained court may never overrule a precedent rule. On other versions, the constraint may be much weaker. On no version, however, is a constrained court as free to disregard a precedent rule as it is on the natural model. The constrained court cannot decide to overrule merely because, having weighed equality and reliance against the advantages of a different rule or decision, it has found the balance slightly tilted in favor of the latter.

The second dimension along which versions of the rule model vary is in their methodologies for identifying the precedent rule. Some versions try to locate a statement of a canonical rule in the opinion of the precedent court. Other versions look for rules that the precedent court implicitly as well as explicitly used as necessary steps in reaching its decision or facts that it deemed material to its decision. On all versions of the rule model, however, the identification of the rule must meet three conditions. First, the rule must have a canonical formulation, even if that canonical formulation does not appear in the original opinion, such as, "Whenever facts, A, B, and C, and not fact D, decide for P." Second, the rule must be treated as separate from the reasoning that led to its adoption by the precedent court; it is only the rule, and not the reasoning, that binds the constrained court. Third, the formulation of the rule must be fixed at the time of the precedent decision; that is, it must not be dependent on what any court other than the precedent court did.

As long as these three conditions are met, coupled with the condition that the constraint be greater than under the natural model of precedent, we are dealing with the rule model. Thus, issues such as whether an announced rule can be a precedent when the precedent court (mistakenly) reached a decision opposite from that which its rule dictates or, relatedly, whether rules intended to apply prospectively only can be binding are issues that arise within the rule model of precedent. The resolution of these questions as well as the resolution of the general debate over how to identify the precedent rule do not affect the general points that I wish to about the rule model. If the model is otherwise attractive, however, their resolution will become necessary.

B. Advantages of the Rule Model of Precedent

The rule model of precedent has several advantages over the natural model. If one considers equality to be a value supporting precedential constraint . . . then the rule model has the advantage of providing access to all the facts of the precedent case relevant to equality. All one needs to know about the precedent case is its rule, which establishes the relevant criteria of equal treatment.

The most important advantage of the rule model, however, is that lower courts (and people in general) derive much more guidance from constraining general rules than they do from constraining particular decisions, even when the particular decisions take the value of rules into account. In claiming this, I am making three assumptions. First, I am assuming a practice of precedent following—precedent here consisting of the rules laid down in precedent cases—that is at least moderately strong. In other words, the

constrained court cannot overrule the precedent merely because it can think of a rule that is slightly superior to the precedent court's rule. The rule model collapses into the natural model if the precedent can be so easily overruled. The advantage that the rule model provides in terms of guidance does not require that precedential constraint be absolute, but it does require that it be greater than zero.

Second, I am assuming that rules provide greater predictability than do other factors to which people might resort in order to predict courts' decisions—factors such as the judges' politics, economic class, and so forth.

Third, I am assuming that an improvement in legal predictability can be a net gain in terms of whatever political morality we hold, even when it comes at the cost of setbacks under that same political morality in particular cases and to particular litigants. This assumption holds up, I believe, despite its paradoxical quality, at least for any political morality that seems at all plausible. . . .

III. THE THIRD MODEL: THE RESULT MODEL OF PRECEDENT

Many readers will feel that I have as yet failed to describe the model of precedential constraint with which they are most familiar. Like Goldilocks and the bowls of porridge and beds, they will complain that the natural model of precedent is too weak to capture their sense of how precedents operate and that the rule model of precedent is too strong. I have not described a model that is "just right."

The third model of precedential constraint appears to meet this complaint. It seems stronger than the natural model in that an incorrect decision has more power to constrain subsequent courts than just reliance or equality values would explain. Yet it also seems weaker than the rule model in that the precedent court's stated rule is not itself binding. . . .

The simplest formulation of the pure result model is as follows: To follow precedent, a constrained court must decide its case for the party analogous to the winner in the precedent case if the constrained case is as strong or stronger a case for that result than the precedent case was for its result. The constrained court must do so even if under the natural model it would have decided its case differently and regardless of any rule stated in the precedent case. Conversely, however, the constrained court may depart from the precedent court's result if the constrained case is a weaker case for that result than was the precedent case, even when the stated rule of the precedent case covers the constrained case and demands a similar result. Put differently, under the pure result model the constrained court must depart from what the natural model requires if and only if its case is an a fortiori case for the same (incorrect) result as in the precedent case.

This formulation of the pure result model of precedent can be presented schematically as follows: Assume that the relevant facts of the precedent case are A, B, C, and D and the precedent court holds for P, declaring as its rule that in all cases of A and B, the decision must be for P. In the constrained case the facts are A, B, C, and E (rather than D) plus N, the natural weight (equality and reliance) of the precedent case. If the absence of D and the presence of E and N make the constrained case a weaker case for P than the precedent case, the constrained court may ignore the rule of the precedent case ("if A and B, decide for P") and decide against P. On the other hand, if the presence of E is a stronger reason for deciding for P than was the presence of D, then even if the constrained court believes that, giving precedent its natural weight, P ought not prevail, the pure result model of precedent constrains it to decide in favor of P. . . .

The Problem of Access to the Facts of the Precedent Case A . . . problem with the pure result model is that it requires the constrained court to have access to those facts of the precedent case that bear on the constrained court's assessment of whether its case is or is not an a fortiori case given the precedent. The precedent court might be quite obliging and recite volumes of facts in its opinion regardless of *its* view of their relevance. On the other hand, the precedent court may believe that its case and future cases should be governed by a very broad rule, which in turn may lead it to recite only those facts pertinent to the rule and only at the level of generality at which the rule operates. For example, it may reveal only that defendant is a promisor and plaintiff a

promisee if it believes that the appropriate rule is that all promises should be legally enforced.

The problem raised by a spare recitation of facts characterized at a high level of generality is that it seriously impedes the constrained court's determination of whether the constrained case is an a fortiori case or is instead distinguishable on its facts from the precedent case. If all the constrained court knows about the precedent case is that the plaintiff was a promisee, the defendant was a promisor, and the plaintiff won, then if it believes its plaintiff-promisor should not win (in the absence of the precedent case), it will be unable to determine whether its case is factually different from the precedent case in a way that supports the defendant (that is, that it is not an a fortiori case for the plaintiff).

If this lack of access to the facts of the precedent case causes the constrained court to follow the precedent decision and to decide, contrary to its natural model judgment, for the party analog to who won in the precedent case, the pure result model of precedent all but collapses into the rule model. At least it does so if the precedent court is clever and reveals only that the facts picked out by its rule obtain. In other words, the advantage that the pure result model purports to have over the rule model—the absence of judicial legislation—is quite fragile and easily defeated by a clever court. . . .

Thus, the pure result model is, because of lack of direct access to the facts of the precedent case, threatened with collapse into either the rule model or the natural model. Neither of these models, however, are hampered by the problem of the precedent court's characterization of the facts of the precedent case. The rule model clearly is not; the natural model merely asks the constrained court to do the best it can with whatever facts it has. On the natural model, limitations on the constrained court's knowledge about the facts of the precedent case are no more significant than other limitations on its information about the world, limitations that are the hallmark of the human condition generally.

REVIEW AND DISCUSSION QUESTIONS

1. How does the rock concert example illustrate the natural model? What two values or rationales support the natural model? Do you agree that courts have some reason to follow prior decisions even if these were incorrect?
2. What is the rule model and what advantages does it have, according to Alexander, over the natural model?
3. Explain the pure result model. How does it differ from the first two models?
4. What problem does Alexander identify with the pure result model?
5. Which of the three models do you think is best? How much weight should courts give to precedent? When and why and to what extent should they follow precedent?

Precedent and Legitimacy

Planned Parenthood v. Casey

Many people expected that after the appointment of Supreme Court Justices O'Connor, Kennedy, and Souter—all of whom were thought to oppose the *Roe v. Wade* decision protecting a woman's right to an abortion—the Court's position would soon change and *Roe* would be overturned. To the surprise of many, however, the three did not reject *Roe* when the opportunity arose but instead wrote a joint opinion explaining that while they may disagree with the decision on the merits, it would nonetheless be a mistake to overturn it twenty years later. Justice Scalia wrote a vigorous dissent.

112 S.Ct. 2791 (1992).

JUSTICES O'CONNOR, KENNEDY, and SOUTER: While we appreciate the weight of the arguments made on behalf of the State in the case before us, arguments which in their ultimate formulation conclude that Roe should be overruled, the reservations any of us may have in reaffirming the central holding of Roe are outweighed by the explication of individual liberty we have given combined with the force of stare decisis. We turn now to that doctrine.

. . . With Cardozo, we recognize that no judicial system could do society's work if it eyed each issue afresh in every case that raised it. *The Nature of the Judicial Process* 149 (1921). Indeed, the very concept of the rule of law [requires] such continuity over time that a respect for precedent is, by definition, indispensable. At the other extreme, a different necessity would make itself felt if a prior judicial ruling should come to be seen so clearly as error that its enforcement was for that very reason doomed.

Even when the decision to overrule a prior case is not [virtually] foreordained, [the] rule of stare decisis is not an "inexorable command," and certainly it is not such in every constitutional case. . . . Rather, when this Court reexamines a prior holding, its judgment is customarily informed by a series of prudential and pragmatic considerations designed to test the consistency of overruling a prior decision with the ideal of the rule of law, and to gauge the respective costs of reaffirming and overruling a prior case. . . .

So in this case we may inquire whether Roe's central rule has been found unworkable; whether the rule's limitation on state power could be removed without serious inequity to those who have relied upon it or significant damage to the stability of the society governed by the rule in question; whether the law's growth in the intervening years has left Roe's central rule a doctrinal anachronism discounted by society; and whether Roe's premises of fact have so far changed in the ensuing two decades as to render its central holding somehow irrelevant or unjustifiable in dealing with the issue it addressed.

1. Although Roe has engendered opposition, it has in no sense proven "unworkable," representing as it does a simple limitation beyond which a state law is unenforceable. While Roe has [required] judicial assessment of [laws] affecting the exercise of the choice guaranteed against government infringement, and although the need for such review will remain as a consequence of today's decision, [these] determinations fall within judicial competence.

2. The inquiry into reliance counts the cost of a rule's repudiation as it would fall on those who have relied reasonably on the rule's continued application. . . . For two decades of economic and social developments, people have organized intimate relationships and made choices that define their views of themselves and their places in society, in reliance on the availability of abortion in the event that contraception should fail. The ability of women to participate equally in the economic and social life of the Nation has been facilitated by their ability to control their reproductive lives. [While] the effect of reliance on Roe cannot be exactly measured, neither can the certain cost of overruling Roe for people who have ordered their thinking and living around that case be dismissed.

3. No evolution of legal principle has left Roe's doctrinal footings weaker than they were in 1973. . . . Roe, however, may be seen not only as an exemplar of Griswold liberty but as a rule (whether or not mistaken) of personal autonomy and bodily integrity, with doctrinal affinity to cases recognizing limits on governmental power to mandate medical treatment or to bar its rejection. If so, our cases since Roe accord with Roe's view that a State's interest in the protection of life falls short of justifying any plenary override of individual liberty claims.

. . . Even on the assumption that the central holding of Roe was in error, that error would go only to the strength of the state interest in fetal protection, not to the recognition afforded woman's liberty. . . . If indeed the woman's interest in deciding whether to bear and beget a child had not been recognized as in Roe, the State might as readily restrict a woman's right to choose to carry a pregnancy to term as to terminate it, to further asserted state interests in population control, or eugenics, for example. Yet Roe has been sensibly relied upon to counter any such suggestions.

4. [Time] has overtaken some of Roe's factual assumptions: advances in maternal health care

allow for abortions safe to the mother later in pregnancy than was true in 1973, and advances in neonatal care have advanced viability to a point somewhat earlier. But these facts . . . have no bearing on the central holding that viability marks the earliest point at which the State's interest in fetal life is constitutionally adequate to justify a legislative ban on nontherapeutic abortions. The soundness or unsoundness of that constitutional judgment in no sense turns on whether viability occurs at approximately 28 weeks, as was usual at the time of Roe; at 25 to 24 weeks, as it sometimes does today. . . . Whenever it may occur, the attainment of viability may continue to serve as the critical fact, just as it has done since Roe; [no] change in Roe's factual underpinning has left its central holding obsolete, and none supports an argument for overruling it.

5. The sum of the precedential inquiry to this point shows Roe's underpinnings unweakened in any way affecting its central holding. While it has engendered disapproval, it has not been unworkable. An entire generation has come of age free to assume Roe's concept of liberty in defining the capacity of women to act in society, and to make reproductive decisions; no erosion of principle going to liberty or personal autonomy has left Roe's central holding a doctrinal remnant; Roe portends no developments at odds with other precedent for the analysis of personal liberty; and no changes of fact have rendered viability more or less appropriate as the point at which the balance of interests tips. Within the bounds of normal stare decisis analysis, [the] stronger argument is for affirming Roe's central holding, with whatever degree of personal reluctance any of us may have, not for overruling it. . . .

. . . [As] Americans of each succeeding generation are rightly told, the Court cannot buy support for its decisions by spending money and, except to a minor degree, it cannot independently coerce obedience to its decrees. The Court's power lies, rather, in its legitimacy, a product of substance and perception that shows itself in the people's acceptance of the Judiciary as fit to determine what the Nation's law means and to declare what it demands. . . . Because not every conscientious claim of principled justification will be accepted as such, the justification claimed must be beyond dispute. The Court must take care to speak and act in ways that allow people to accept its decisions on the terms the Court claims for them, as grounded truly in principle, not as compromises with social and political pressures having, as such, no bearing on the principled choices that the Court is obliged to make. Thus, the Court's legitimacy depends on making legally principled decisions under circumstances in which their principled character is sufficiently plausible to be accepted by the Nation.

. . . In two circumstances, however, the Court would almost certainly fail to receive the benefit of the doubt in overruling prior cases. There is, first, a point beyond which frequent overruling would overtax the country's belief in the Court's good faith. . . . There is a limit to the amount of error that can plausibly be imputed to prior courts. If that limit should be exceeded, disturbance of prior rulings would be taken as evidence that justifiable reexamination of principle had given way to drives for particular results in the short term. The legitimacy of the Court would fade with the frequency of its vacillation.

That first circumstance can be described as hypothetical; the second is to the point here and now. Where, in the performance of its judicial duties, the Court decides a case in such a way as to resolve the sort of intensely divisive controversy reflected in Roe and those rare, comparable cases, its decision has a dimension that the resolution of the normal case does not carry. . . . So to overrule under fire in the absence of the most compelling reason to reexamine a watershed decision would subvert the Court's legitimacy beyond any serious question. . . . An extra price will be paid by those who themselves disapprove of the decision's results when viewed outside of constitutional terms, but who nevertheless struggle to accept it, because they respect the rule of law. To all those who will be so tested by following, the Court implicitly undertakes to remain steadfast, lest in the end a price be paid for nothing. The promise of constancy, once given, binds its maker for as long as the power to stand by the decision survives and the understanding of the issue has not changed so fundamentally as to render the commitment obsolete.

[Like] the character of an individual, the legitimacy of the Court must be earned over time. So,

indeed, must be the character of a Nation of people who aspire to live according to the rule of law. Their belief in themselves as such a people is not readily separable from their understanding of the Court invested with the authority to decide their constitutional cases and speak before all others for their constitutional ideals. If the Court's legitimacy should be undermined, then, so would the country be in its very ability to see itself through its constitutional ideals. The Court's concern with legitimacy is not for the sake of the Court but for the sake of the Nation.

. . . Whether or not a new social consensus is developing on that issue, its divisiveness is no less today than in 1973, and pressure to overrule the decision, like pressure to retain it, has grown only more intense. A decision to overrule Roe's essential holding under the existing circumstances would address error, if error there was, at the cost of [profound] and unnecessary damage to the Court's legitimacy, and to the Nation's commitment to the rule of law. It is therefore imperative to adhere to the essence of Roe's original decision, and we do so today. . . .

Justice Scalia [dissenting]: . . . I would uphold the Pennsylvania statute in its entirety. . . . Assuming that the question before us is to be resolved at such a level of philosophical abstraction, in such isolation from the traditions of American society, as by simply applying "reasoned judgment," I do not see how that could possibly have produced the answer the Court arrived at in Roe. Today's opinion describes the methodology of Roe, quite accurately, as weighing against the woman's interest the State's "important and legitimate interest in protecting the potentiality of human life." But "reasoned judgment" does not begin by begging the question, as Roe unquestionably did by assuming that what the State is protecting is the mere "potentiality of human life." The whole argument of abortion opponents is that what the Court calls the fetus and what others call the unborn child is a human life. Thus, whatever answer Roe came up with after conducting its "balancing" is bound to be wrong, unless it is correct that the human fetus is in some critical sense merely potentially human. There is of course no way to determine that as a legal matter; it is in fact a value judgment. Some societies have considered newborn children not yet human, or the incompetent elderly no longer so.

The authors of the joint opinion [do] not squarely contend that Roe was a correct application of "reasoned judgment," merely that it must be followed, because of stare decisis. But in their exhaustive discussion of all the factors that go into the determination of when stare decisis should be observed and when disregarded, they never mention "how wrong was the decision on its face?" Surely, if "the Court's power lies . . . in its legitimacy, a product of substance and perception," the "substance" part of the equation demands that plain error be acknowledged and eliminated. Roe was plainly wrong—even on the Court's methodology of "reasoned judgment," and even more so (of course) if the proper criteria of text and tradition are applied. The emptiness of the "reasoned judgment" that produced Roe is displayed in plain view by the fact that, after more than 19 years of effort by some of the brightest (and most determined) legal minds in the country, after more than 10 cases upholding abortion rights in this Court, and after dozens upon dozens of amicus briefs submitted in this and other cases, the best the Court can do to explain how it is that the word "liberty" must be thought to include the right to destroy human fetuses is to rattle off a collection of adjectives that simply decorate a value judgment and conceal a political choice. The right to abort, we are told, inheres in "liberty" because it is among "a person's most basic decisions"; it involves a "most intimate and personal choic[e]"; it is "central to personal dignity and autonomy"; it "originate[s] within the zone of conscience and belief"; it is "too intimate and personal" for state interference; it reflects "intimate views" of a "deep, personal character"; it involves "intimate relationships," and notions of "personal autonomy and bodily integrity"; and it concerns a particularly "important decision." But it is obvious to anyone applying "reasoned judgment" that the same adjectives can be applied to many forms of conduct that this Court has held are not entitled to constitutional protection—because, like abortion, they are forms of conduct that have long been criminalized in American society. Those adjectives might be applied, for example, to homosexual sodomy, polygamy, adult incest, and suicide, all of which are equally "intimate" and "deeply

personal" decisions involving "personal autonomy and bodily integrity," and all of which can constitutionally be proscribed because it is our unquestionable constitutional tradition that they are proscribable. It is not reasoned judgment that supports the Court's decision; only personal predilection. . . . It is particularly difficult, in the circumstances of the present decision, to sit still for the Court's lengthy lecture upon the virtues of "constancy," of "remaining steadfast," of adhering to "principle." Among the five Justices who purportedly adhere to Roe, at most three agree upon the principle that constitutes adherence (the joint opinion's "undue burden" standard)—and that principle is inconsistent with Roe. To make matters worse, two of the three, in order thus to remain steadfast, had to abandon previously stated positions. It is beyond me how the Court expects these accommodations to be accepted "as grounded truly in principle, not as compromises with social and political pressures having, as such, no bearing on the principled choices that the Court is obliged to make." The only principle the Court "adheres" to, it seems to me, is the principle that the Court must be seen as standing by Roe. That is not a principle of law (which is what I thought the Court was talking about), but a principle of Realpolitik—and a wrong one at that.

I cannot agree with, indeed I am appalled by, the Court's suggestion that the decision whether to stand by an erroneous constitutional decision must be strongly influenced—against overruling, no less—by the substantial and continuing public opposition the decision has generated. The Court's judgment that any other course would "subvert the Court's legitimacy" must be another consequence of reading the error-filled history book that described the deeply divided country brought together by Roe. . . .

But whether it would "subvert the Court's legitimacy" or not, the notion that we would decide a case differently from the way we otherwise would have in order to show that we can stand firm against public disapproval is frightening. It is a bad enough idea, even in the head of someone like me, who believes that the text of the Constitution, and our traditions, say what they say and there is no fiddling with them. But when it is in the mind of a Court that believes the Constitution has an evolving meaning . . . and that the function of this Court is to "speak before all others for [the people's] constitutional ideals" unrestrained by meaningful text or tradition—then the notion that the court must adhere to a decision for as long as the decision faces "great opposition" and the Court is "under fire" acquires a character of almost czarist arrogance. . . .

In truth, I am as distressed as the Court is—and expressed my distress several years ago . . . —about the "political pressure" directed to the Court: the marches, the mail, the protests aimed at inducing us to change our opinions. How upsetting it is, that so many of our citizens (good people, not lawless ones, on both sides of this abortion issue, and on various sides of other issues as well) think that we Justices should properly take into account their views, as though we were engaged not in ascertaining an objective law but in determining some kind of social consensus. The Court would profit, I think, from giving less attention to the fact of this distressing phenomenon, and more attention to the cause of it. That cause permeates today's opinion: a new mode of constitutional adjudication that relies not upon text and traditional practice to determine the law, but upon what the Court calls "reasoned judgment," which turns out to be nothing but philosophical predilection and moral intuition. All manner of "liberties," the Court tells us, inhere in the Constitution and are enforceable by this Court—not just those mentioned in the text or established in the traditions of our society. . . .

What makes all this relevant to the bothersome application of "political pressure" against the Court are the twin facts that the American people love democracy and the American people are not fools. As long as this Court thought (and the people thought) that we Justices were doing essentially lawyers' work up here—reading text and discerning our society's traditional understanding of that text—the public pretty much left us alone. Texts and traditions are facts to study, not convictions to demonstrate about. But if in reality our process of constitutional adjudication consists primarily of making value judgments; if we can ignore a long and clear tradition clarifying an ambiguous text, [if our] pronouncement of constitutional law rests primarily on value judgments, then a free and intelligent people's attitude towards us can be expected

to be (ought to be) quite different. The people know that their value judgments are quite as good as those taught in any law school—maybe better. If [the] "liberties" protected by the Constitution are [undefined] and unbounded, then the people should demonstrate, to protest that we do not implement their values instead of ours. Not only that, but confirmation hearings for new Justices should deteriorate into question-and-answer sessions in which Senators go through a list of their constituents' most favored and most disfavored alleged constitutional rights, and seek the nominee's commitment to support or oppose them. Value judgments, after all, should be voted on, not dictated; and if our Constitution has somehow accidentally committed them to the Supreme Court, at least we can have a sort of plebiscite each time a new nominee to that body is put forward.

REVIEW AND DISCUSSION QUESTIONS

1. Explain the reasons why the three justices decided not to overturn *Roe v. Wade.*
2. Justice Scalia argues that a mistake of the past should be corrected, and that the arguments of the three justices will have the opposite effect, perpetuating a mistaken law. Explain his reasoning.
3. Situating your discussion in the context of material you have read on precedent, legal reasoning and the rule of law, discuss whether or not you think this case was correctly decided. Suppose you believed strongly that *Roe* was a mistake at the time; would that affect your conclusion? Explain.

CHAPTER 6

Constitutional Interpretation

The essays in this section focus on one of the most difficult and controversial areas of legal reasoning, namely, constitutional interpretation. In particular, legal philosopher Ronald Dworkin, Supreme Court Justice Antonin Scalia, and federal appellate judge Richard A. Posner provide rival ways of thinking about constitutional issues. And the cases *Rochin v. California* and *District of Columbia v. Heller* illustrate well the frequently controversial character of constitutional adjudication.

"Vague" Constitutional Clauses

Ronald Dworkin

In this extract from an essay in his book *Taking Rights Seriously*, Ronald Dworkin explains that the Constitution upholds certain moral rights that individuals possess against the majority. Some of these rights are stated in language that appears vague and is often controversial, compared to other, more precise parts of the Constitution. Sometimes there is said to be a "strict" and a "liberal" side to these controversies, but Dworkin argues that this way of looking at the matter runs together two different issues. He goes on to distinguish *concepts* and *conceptions* and to argue that fidelity to the text of the Constitution requires courts to decide between competing conceptions of what political morality requires.

The constitutional theory on which our government rests is not a simple majoritarian theory. The Constitution, and particularly the Bill of Rights, is designed to protect individual citizens and groups against certain decisions that a majority of citizens might want to make, even when that majority acts in what it takes to be the general or common interest. Some of these constitutional restraints take the form of fairly precise rules, like the rule that requires a jury trial in federal criminal proceedings or, perhaps, the rule that forbids the national Congress to abridge freedom of speech. But other constraints take the form of what are often called "vague" standards, for example, the provision that

the government shall not deny men due process of law, or equal protection of the laws.

This interference with democratic practice requires a justification. The draftsmen of the Constitution assumed that these restraints could be justified by appeal to moral rights which individuals possess against the majority, and which the constitutional provisions, both "vague" and precise, might be said to recognize and protect.

The "vague" standards were chosen deliberately, by the men who drafted and adopted them, in place of the more specific and limited rules that they might have enacted. But their decision to use the language they did has caused a great deal of legal and political controversy, because even reasonable men of good will differ when they try to elaborate, for example, the moral rights that the due process clause or the equal protection clause brings into the law. They also differ when they try to apply these rights, however defined, to complex matters of political administration, like the educational practices that were the subject of the segregation cases.

The practice has developed of referring to a "strict" and a "liberal" side to these controversies, so that the Supreme Court might be said to have taken the "liberal" side in the segregation cases and its critics the "strict" side. . . . But the distinction is in fact confusing, because it runs together two different issues that must be separated. Any case that arises under the "vague" constitutional guarantees can be seen as posing two questions: (1) Which decision is required by strict, that is to say faithful, adherence to the text of the Constitution or to the intention of those who adopted that text? (2) Which decision is required by a political philosophy that takes a strict, that is to say narrow, view of the moral rights that individuals have against society? Once these questions are distinguished, it is plain that they may have different answers. The text of the First Amendment, for example, says that Congress shall make *no* law abridging the freedom of speech, but a narrow view of individual rights would permit many such laws, ranging from libel and obscenity laws to the Smith Act.

In the case of the "vague" provisions, however, like the due process and equal protection clauses, lawyers have run the two questions together because they have relied, largely without recognizing it, on a theory of meaning that might be put this way: If the framers of the Constitution used vague language, as they did when they condemned violations of "due process of law," then what they "said" or "meant" is limited to the instances of official action that they had in mind as violations, or, at least, to those instances that they would have thought were violations if they had had them in mind. If those who were responsible for adding the due process clause to the Constitution believed that it was fundamentally unjust to provide separate education for different races, or had detailed views about justice that entailed that conclusion, then the segregation decisions might be defended as an application of the principle they had laid down. Otherwise, they could not be defended in this way, but instead would show that the judges had substituted their own ideas of justice for those the constitutional drafters meant to lay down.

This theory makes a strict interpretation of the text yield a narrow view of constitutional rights, because it limits such rights to those recognized by a limited group of people at a fixed date of history. It forces those who favor a more liberal set of rights to concede that they are departing from strict legal authority, a departure they must then seek to justify by appealing only to the desirability of the results they reach.

But the theory of meaning on which this argument depends is far too crude; it ignores a distinction that philosophers have made but lawyers have not yet appreciated. Suppose I tell my children simply that I expect them not to treat others unfairly. I no doubt have in mind examples of the conduct I mean to discourage, but I would not accept that my "meaning" was limited to these examples, for two reasons. First, I would expect my children to apply my instructions to situations I had not and could not have thought about. Second, I stand ready to admit that some particular act I had thought was fair when I spoke was in fact unfair, or vice versa, if one of my children is able to convince me of that later; in that case I should want to say that my instructions covered the case he cited, not that I had changed my instructions. I might say that I meant the family to be guided by the *concept* of fairness, not by any specific *conceptions* of fairness I might have had in mind.

This is a crucial distinction which it is worth pausing to explore. Suppose a group believes in common that acts may suffer from a special moral defect which they call unfairness, and which consists in a wrongful division of benefits and burdens, or a wrongful attribution of praise or blame. Suppose also that they agree on a great number of standard cases of unfairness and use these as benchmarks against which to test other, more controversial cases. In that case, the group has a concept of unfairness, and its members may appeal to that concept in moral instruction or argument. But members of that group may nevertheless differ over a large number of these controversial cases, in a way that suggests that each either has or acts on a different theory of *why* the standard cases are acts of unfairness. They may differ, that is, on which more fundamental principles must be relied upon to show that a particular division or attribution is unfair. In that case, the members have different conceptions of fairness.

If so, then members of this community who give instructions or set standards in the name of fairness may be doing two different things. First they may be appealing to the concept of fairness, simply by instructing others to act fairly; in this case they charge those whom they instruct with the responsibility of developing and applying their own conception of fairness as controversial cases arise. That is not the same thing, of course, as granting them a discretion to act as they like; it sets a standard which they must try—and may fail—to meet, because it assumes that one conception is superior to another. The man who appeals to the concept in this way may have his own conception, as I did when I told my children to act fairly; but he holds this conception only as his own theory of how the standard he set must be met, so that when he changes his theory he has not changed that standard.

On the other hand, the members may be laying down a particular conception of fairness; I would have done this, for example, if I had listed my wishes with respect to controversial examples or if, even less likely, I had specified some controversial and explicit theory of fairness, as if I had said to decide hard cases by applying the utilitarian ethics of Jeremy Bentham. The difference is a difference not just in the *detail* of the instructions given but in the *kind* of instructions given. When I appeal to the concept of fairness I appeal to what fairness means, and I give my views on that issue no special standing. When I lay down a conception of fairness, I lay down what I mean by fairness, and my view is therefore the heart of the matter. When I appeal to fairness I pose a moral issue; when I lay down my conception of fairness I try to answer it.

Once this distinction is made it seems obvious that we must take what I have been calling "vague" constitutional clauses as representing appeals to the concepts they employ, like legality, equality, and cruelty. The Supreme Court may soon decide, for example, whether capital punishment is "cruel" within the meaning of the constitutional clause that prohibits "cruel and unusual punishment." It would be a mistake for the Court to be much influenced by the fact that when the clause was adopted capital punishment was standard and unquestioned. That would be decisive if the framers of the clause had meant to lay down a particular conception of cruelty, because it would show that the conception did not extend so far. But it is not decisive of the different question the Court now faces, which is this: Can the Court, responding to the framers' appeal to the concept of cruelty, now defend a conception that does not make death cruel?

Those who ignore the distinction between concepts and conceptions, but who believe that the Court ought to make a fresh determination of whether the death penalty is cruel, are forced to argue in a vulnerable way. They say that ideas of cruelty change over time, and that the Court must be free to reject out-of-date conceptions; this suggests that the Court must change what the Constitution enacted. But in fact the Court can enforce what the Constitution says only by making up its own mind about what is cruel, just as my children, in my example, can do what I said only by making up their own minds about what is fair. If those who enacted the broad clauses had meant to lay down particular conceptions, they would have found the sort of language conventionally used to do this, that is, they would have offered particular theories of the concepts in question.

Indeed the very practice of calling these clauses "vague," in which I have joined, can now be seen to involve a mistake. The clauses are vague only if

we take them to be botched or incomplete or schematic attempts to lay down particular conceptions. If we take them as appeals to moral concepts they could not be made more precise by being more detailed.[1]

The confusion I mentioned between the two senses of "strict" construction is therefore very misleading indeed. If courts try to be faithful to the text of the Constitution, they will for that very reason be forced to decide between competing conceptions of political morality. So it is wrong to attack the Warren Court, for example, on the ground that it failed to treat the Constitution as a binding text. On the contrary, if we wish to treat fidelity to that text as an overriding requirement of constitutional interpretation, then it is the conservative critics of the Warren Court who are at fault, because their philosophy ignores the direction to face issues of moral principle that the logic of the text demands.

I put the matter in a guarded way because we may *not* want to accept fidelity to the spirit of the text as an overriding principle of constitutional adjudication. It may be more important for courts to decide constitutional cases in a manner that respects that judgments of other institutions of government, for example. Or it may be more important for courts to protect established legal doctrines, so that citizens and the government can have confidence that the courts will hold to what they have said before. But it is crucial to recognize that these other policies compete with the principle that the Constitution is the fundamental and imperative source of constitutional law. They are not, as the "strict constructionists" suppose, simply consequences of that principle.

NOTE

1. It is less misleading to say that the broad clauses of the Constitution "delegate" power to the Court to enforce its own conceptions of political morality. But even this is inaccurate if it suggests that the Court need not justify its conception by arguments showing the connections between its conception and standard cases, as described in the text. If the Court finds that the death penalty is cruel, it must do so on the basis of some principles or groups of principles that unite the death penalty with the thumbscrew and the rack.

REVIEW AND DISCUSSION QUESTIONS

1. Give examples of precise constitutional rules and of vague constitutional standards.
2. Dworkin distinguishes two senses of "strict" construction but accepts only one of them. What are the two?
3. Describe the distinction between a concept and a conception. How is it illustrated by the parent telling his or her children to be fair?
4. What does this distinction imply for constitutional interpretation? Why would it be wrong for the Court to answer the question whether capital punishment is "cruel" based on what the framers of the Constitution thought?

The Due Process Clause

Rochin v. California

Writing for the majority in this famous Supreme Court case, Justice Frankfurter ruled that police treatment of defendant Rochin violated the Fourteenth Amendment's requirement that no state "deprive any person of life, liberty, or property, without due process of law." Although the states have primary responsibility for defining and enforcing the criminal law, the Court has a duty to ensure

that criminal proceedings do not offend the "canons of decency and fairness." The requirements of the Due Process Clause are not fixed and exact, but indefinite and vague; however, this is not unusual when dealing with human rights, Frankfurter argues, and judges can exercise reasoned judgment to determine when state conduct violates basic constitutional principles. In their concurring opinions, Justices Black and Douglas argue that the majority's use of the Due Process Clause is too subjective and that Rochin should instead be acquitted on Fifth Amendment grounds. (In his dissent in an earlier case, *Adamson v. California*, Black had argued that the Fourteenth Amendment requires states to guarantee defendants the full protection of the Fifth Amendment—a position that the Court rejected then but was to adopt, in effect, in the years after *Rochin*).

Mr. Justice Frankfurter delivered the opinion of the Court: Having "some information that [the petitioner here] was selling narcotics," three deputy sheriffs of the County of Los Angeles, on the morning of July 1, 1949, made for the two-story dwelling house in which Rochin lived with his mother, common-law wife, brothers and sisters. Finding the outside door open, they entered and then forced open the door to Rochin's room on the second floor. Inside they found petitioner sitting partly dressed on the side of the bed, upon which his wife was lying. On a "night stand" beside the bed the deputies spied two capsules. When asked "Whose stuff is this?" Rochin seized the capsules and put them in his mouth. A struggle ensued, in the course of which the three officers "jumped upon him" and attempted to extract the capsules. The force they applied proved unavailing against Rochin's resistance. He was handcuffed and taken to a hospital. At the direction of one of the officers a doctor forced an emetic solution through a tube into Rochin's stomach against his will. This "stomach pumping" produced vomiting. In the vomited matter were found two capsules which proved to contain morphine.

Rochin was brought to trial before a California Superior Court, sitting without a jury, on the charge of possessing "a preparation of morphine" in violation of the California Health and Safety Code, 1947, 11,500. Rochin was convicted and sentenced to sixty days' imprisonment. The chief evidence against him was the two capsules. They were admitted over petitioner's objection, although the means of obtaining them was frankly set forth in the testimony by one of the deputies, substantially as here narrated.

On appeal, the District Court of Appeal affirmed the conviction, despite the finding that the officers "were guilty of unlawfully breaking into and entering defendant's room and were guilty of unlawfully assaulting and battering defendant while in the room," and "were guilty of unlawfully assaulting, battering, torturing and falsely imprisoning the defendant at the alleged hospital." One of the three judges, while finding that "the record in this case reveals a shocking series of violations of constitutional rights," concurred only because he felt bound by decisions of his Supreme Court. These, he asserted, "have been looked upon by law enforcement officers as an encouragement, if not an invitation, to the commission of such lawless acts." The Supreme Court of California denied without opinion Rochin's petition for a hearing. Two justices dissented from this denial, and in doing so expressed themselves thus: " . . . a conviction which rests upon evidence of incriminating objects obtained from the body of the accused by physical abuse is as invalid as a conviction which rests upon a verbal confession extracted from him by such abuse. . . . Had the evidence forced from the defendant's lips consisted of an oral confession that he illegally possessed a drug . . . he would have the protection of the rule of law which excludes coerced confessions from evidence. But because the evidence forced from his lips consisted of real objects the People of this state are permitted to base a conviction upon it. [We] find no valid ground of distinction between a verbal confession extracted by physical abuse and a confession wrested from defendant's body by physical abuse."

This Court granted certiorari because a serious question is raised as to the limitations which the Due Process Clause of the Fourteenth Amendment imposes on the conduct of criminal proceedings by the States.

342 U.S. 165 (1952)

In our federal system the administration of criminal justice is predominantly committed to the care of the States. The power to define crimes belongs to Congress only as an appropriate means of carrying into execution its limited grant of legislative powers. Broadly speaking, crimes in the United States are what the laws of the individual States make them, subject to [certain] limitations. . . .

These limitations, in the main, concern not restrictions upon the powers of the States to define crime, except in the restricted area where federal authority has preempted the field, but restrictions upon the manner in which the States may enforce their penal codes. Accordingly, in reviewing a State criminal conviction under a claim of right guaranteed by the Due Process Clause of the Fourteenth Amendment, from which is derived the most far-reaching and most frequent federal basis of challenging State criminal justice, "we must be deeply mindful of the responsibilities of the States for the enforcement of criminal laws, and exercise with due humility our merely negative function in subjecting convictions from state courts to the very narrow scrutiny which the Due Process Clause of the Fourteenth Amendment authorizes." *Malinski v. New York.* Due process of law, "itself a historical product," is not to be turned into a destructive dogma against the States in the administration of their systems of criminal justice.

However, this Court too has its responsibility. Regard for the requirements of the Due Process Clause "inescapably imposes upon this Court an exercise of judgment upon the whole course of the proceedings [resulting in a conviction] in order to ascertain whether they offend those canons of decency and fairness which express the notions of justice of English-speaking peoples even toward those charged with the most heinous offenses." *Malinski v. New York.* These standards of justice are not authoritatively formulated anywhere as though they were specifics. Due process of law is a summarized constitutional guarantee of respect for those personal immunities which, as Mr. Justice Cardozo twice wrote for the Court, are "so rooted in the traditions and conscience of our people as to be ranked as fundamental" or are "implicit in the concept of ordered liberty."

The Court's function in the observance of this settled conception of the Due Process Clause does not leave us without adequate guides in subjecting State criminal procedures to constitutional judgment. In dealing not with the machinery of government but with human rights, the absence of formal exactitude, or want of fixity of meaning, is not an unusual or even regrettable attribute of constitutional provisions. Words being symbols do not speak without a gloss. On the one hand the gloss may be the deposit of history, whereby a term gains technical content. Thus the requirements of the Sixth and Seventh Amendments for trial by jury in the federal courts have a rigid meaning. No changes or chances can alter the content of the verbal symbol of "jury"—a body of twelve men who must reach a unanimous conclusion if the verdict is to go against the defendant. On the other hand, the gloss of some of the verbal symbols of the Constitution does not give them a fixed technical content. It exacts a continuing process of application.

When the gloss has thus not been fixed but is a function of the process of judgment, the judgment is bound to fall differently at different times and differently at the same time through different judges. Even more specific provisions, such as the guaranty of freedom of speech and the detailed protection against unreasonable searches and seizures, have inevitably evoked as sharp divisions in this Court as the least specific and most comprehensive protection of liberties, the Due Process Clause.

The vague contours of the Due Process Clause do not leave judges at large. We may not draw on our merely personal and private notions and disregard the limits that bind judges in their judicial function. Even though the concept of due process of law is not final and fixed, these limits are derived from considerations that are fused in the whole nature of our judicial process. These are considerations deeply rooted in reason and in the compelling traditions of the legal profession. The Due Process Clause places upon this Court the duty of exercising a judgment, within the narrow confines of judicial power in reviewing State convictions, upon interests of society pushing in opposite directions.

Due process of law thus conceived is not to be derided as resort to a revival of "natural law." To believe that this judicial exercise of judgment could be avoided by freezing "due process of law"

at some fixed stage of time or thought is to suggest that the most important aspect of constitutional adjudication is a function for inanimate machines and not for judges. . . . To practice the requisite detachment and to achieve sufficient objectivity no doubt demands of judges the habit of self-discipline and self-criticism, incertitude that one's own views are incontestable and alert tolerance toward views not shared. But these are precisely the presuppositions of our judicial process. They are precisely the qualities society has a right to expect from those entrusted with ultimate judicial power.

Restraints on our jurisdiction are self-imposed only in the sense that there is from our decisions no immediate appeal short of impeachment or constitutional amendment. But that does not make due process of law a matter of judicial caprice. The faculties of the Due Process Clause may be indefinite and vague, but the mode of their ascertainment is not self-willed. In each case "due process of law" requires an evaluation based on a disinterested inquiry pursued in the spirit of science, on a balanced order of facts exactly and fairly stated, on the detached consideration of conflicting claims, on a judgment not ad hoc and episodic but duly mindful of reconciling the needs both of continuity and of change in a progressive society.

Applying these general considerations to the circumstances of the present case, we are compelled to conclude that the proceedings by which this conviction was obtained do more than offend some fastidious squeamishness or private sentimentalism about combatting crime too energetically. This is conduct that shocks the conscience. Illegally breaking into the privacy of the petitioner, the struggle to open his mouth and remove what was there, the forcible extraction of his stomach's contents—this course of proceeding by agents of government to obtain evidence is bound to offend even hardened sensibilities. They are methods too close to the rack and the screw to permit of constitutional differentiation.

It has long since ceased to be true that due process of law is heedless of the means by which otherwise relevant and credible evidence is obtained. This was not true even before the series of recent cases enforced the constitutional principle that the States may not base convictions upon confessions, however much verified, obtained by coercion. These decisions are not arbitrary exceptions to the comprehensive right of States to fashion their own rules of evidence for criminal trials. They are not sports in our constitutional law but applications of a general principle. They are only instances of the general requirement that States in their prosecutions respect certain decencies of civilized conduct. Due process of law, as a historic and generative principle, precludes defining, and thereby confining, these standards of conduct more precisely than to say that convictions cannot be brought about by methods that offend "a sense of justice." It would be a stultification of the responsibility which the course of constitutional history has cast upon this Court to hold that in order to convict a man the police cannot extract by force what is in his mind but can extract what is in his stomach.

To attempt in this case to distinguish what lawyers call "real evidence" from verbal evidence is to ignore the reasons for excluding coerced confessions. Use of involuntary verbal confessions in State criminal trials is constitutionally obnoxious not only because of their unreliability. They are inadmissible under the Due Process Clause even though statements contained in them may be independently established as true. Coerced confessions offend the community's sense of fair play and decency. So here, to sanction the brutal conduct which naturally enough was condemned by the court whose judgment is before us, would be to afford brutality the cloak of law. Nothing would be more calculated to discredit law and thereby to brutalize the temper of a society. . . .

We are not unmindful that hypothetical situations can be conjured up, shading imperceptibly from the circumstances of this case and by gradations producing practical differences despite seemingly logical extensions. But the Constitution is "intended to preserve practical and substantial rights, not to maintain theories." *Davis v. Mills.*

On the facts of this case the conviction of the petitioner has been obtained by methods that offend the Due Process Clause. The judgment below must be

Reversed.

Mr. Justice Black, concurring: [My dissenting opinion in] *Adamson v. California* sets out reasons for my belief that state as well as federal courts and law enforcement officers must obey the Fifth Amendment's command that "No person . . . shall be compelled in any criminal case to be a witness against himself." I think a person is compelled to be a witness against himself not only when he is compelled to testify, but also when as here, incriminating evidence is forcibly taken from him by a contrivance of modern science. California convicted this petitioner by using against him evidence obtained in this manner, and I agree with MR. JUSTICE DOUGLAS that the case should be reversed on this ground.

In the view of a majority of the Court, however, the Fifth Amendment imposes no restraint of any kind on the states. They nevertheless hold that California's use of this evidence violated the Due Process Clause of the Fourteenth Amendment. Since they hold as I do in this case, I regret my inability to accept their interpretation without protest. But I believe that faithful adherence to the specific guarantees in the Bill of Rights insures a more permanent protection of individual liberty than that which can be afforded by the nebulous standards stated by the majority.

What the majority hold is that the Due Process Clause empowers this Court to nullify any state law if its application "shocks the conscience," offends "a sense of justice" or runs counter to the "decencies of civilized conduct." The majority emphasize that these statements do not refer to their own consciences or to their senses of justice and decency. For we are told that "we may not draw on our merely personal and private notions"; our judgment must be grounded on "considerations deeply rooted in reason and in the compelling traditions of the legal profession." We are further admonished to measure the validity of state practices, not by our reason, or by the traditions of the legal profession, but by "the community's sense of fair play and decency"; by the "traditions and conscience of our people"; or by "those canons of decency and fairness which express the notions of justice of English-speaking peoples." These canons are made necessary, it is said, because of "interests of society pushing in opposite directions."

If the Due Process Clause does vest this Court with such unlimited power to invalidate laws, I am still in doubt as to why we should consider only the notions of English-speaking peoples to determine what are immutable and fundamental principles of justice. Moreover, one may well ask what avenues of investigation are open to discover "canons" of conduct so universally favored that this Court should write them into the Constitution? All we are told is that the discovery must be made by an "evaluation based on a disinterested inquiry pursued in the spirit of science, on a balanced order of facts."

Some constitutional provisions are stated in absolute and unqualified language such, for illustration, as the First Amendment stating that no law shall be passed prohibiting the free exercise of religion or abridging the freedom of speech or press. Other constitutional provisions do require courts to choose between competing policies, such as the Fourth Amendment which, by its terms, necessitates a judicial decision as to what is an "unreasonable" search or seizure. There is, however, no express constitutional language granting judicial power to invalidate every state law of every kind deemed "unreasonable" or contrary to the Court's notion of civilized decencies; yet the constitutional philosophy used by the majority has, in the past, been used to deny a state the right to fix the price of gasoline and even the right to prevent bakers from palming off smaller for larger loaves of bread. These cases, and others, show the extent to which the evanescent standards of the majority's philosophy have been used to nullify state legislative programs passed to suppress evil economic practices. What paralyzing role this same philosophy will play in the future economic affairs of this country is impossible to predict. Of even graver concern, however, is the use of the philosophy to nullify the Bill of Rights. I long ago concluded that the accordion-like qualities of this philosophy must inevitably imperil all the individual liberty safeguards specifically enumerated in the Bill of Rights. Reflection and recent decisions of this Court sanctioning abridgment of the freedom of speech and press have strengthened this conclusion.

Mr. Justice Douglas, concurring: The evidence obtained from this accused's stomach would be admissible in the majority of states where the question has been raised. So far as the reported cases reveal, the only states which would probably exclude the evidence would be Arkansas, Iowa, Michigan, and Missouri. Yet the Court now says that the rule which the majority of the states have fashioned violates the "decencies of civilized conduct." To that I cannot agree. It is a rule formulated by responsible courts with judges as sensitive as we are to the proper standards for law administration.

As an original matter it might be debatable whether the provision in the Fifth Amendment that no person "shall be compelled in any criminal case to be a witness against himself" serves the ends of justice. Not all civilized legal procedures recognize it. But the choice was made by the Framers, a choice which sets a standard for legal trials in this country. The Framers made it a standard of due process for prosecutions by the Federal Government. If it is a requirement of due process for a trial in the federal courthouse, it is impossible for me to say it is not a requirement of due process for a trial in the state courthouse. That was the issue recently surveyed in *Adamson v. California.* The Court rejected the view that compelled testimony should be excluded and held in substance that the accused in a state trial can be forced to testify against himself. I disagree. Of course an accused can be compelled to be present at the trial, to stand, to sit, to turn this way or that, and to try on a cap or a coat. But I think that words taken from his lips, capsules taken from his stomach, blood taken from his veins are all inadmissible provided they are taken from him without his consent. They are inadmissible because of the command of the Fifth Amendment.

That is an unequivocal, definite and workable rule of evidence for state and federal courts. But we cannot in fairness free the state courts from that command and yet excoriate them for flouting the "decencies of civilized conduct" when they admit the evidence. That is to make the rule turn not on the Constitution but on the idiosyncrasies of the judges who sit here.

The damage of the view sponsored by the Court in this case may not be conspicuous here. But it is part of the same philosophy that produced *Betts v. Brady*, denying counsel to an accused in a state trial against the command of the Sixth Amendment, and *Wolf v. Colorado*, allowing evidence obtained as a result of a search and seizure that is illegal under the Fourth Amendment to be introduced in a state trial. It is part of the process of erosion of civil rights of the citizen in recent years.

REVIEW AND DISCUSSION QUESTIONS

1. What are the facts of the case? What was the response of the District Court of Appeal and of the Supreme Court of California to Rochin's appeal?
2. The Court distinguishes constitutional terms that have a fixed technical meaning (like "jury") from others that require "a continuing process of application." Relate this point to Dworkin's discussion of "concepts" and "conceptions" in the previous essay.
3. What analogy does the Court draw with coerced confessions?
4. Do you agree or disagree that what the police officers did "shocks the conscience," "discredit[s] law," and "brutalize[s] the temper of society"? If so, does it follow that their conduct violated due process?
5. Why do Justices Black and Douglas believe that the conduct of the police violated the Fifth Amendment? Why do they think that this approach is better than one that appeals to the Due Process Clause?
6. Does the majority adequately address the concern that its decision is subjective or invokes "natural law"?

Interpreting the Constitution

Antonin Scalia

Supreme Court Justice Antonin Scalia argues that judges should not search for legislative intent but instead focus on what the law says. He calls this general philosophy of interpretation "textualism." Applied to the Constitution, it is the doctrine of orginalism, namely, that the Constitution should be interpreted according to its original meaning, not its current meaning. Accordingly, Scalia rejects the ideal of a Living Constitution that evolves with the times. Because there is no agreement on how the Constitution should evolve or what it ought to say, the Living Constitution approach politicizes the selection of judges and ultimately reduces constitutional interpretation to majority rule.

INTENT

Statutory interpretation is such a broad subject that the substance of it cannot be discussed comprehensively here. It is worth examining a few aspects, however, if only to demonstrate the great degree of confusion that prevails. We can begin at the most fundamental possible level. So utterly unformed is the American law of statutory interpretation that not only is its methodology unclear, but even its very *objective* is. Consider the basic question: What are we looking for when we construe a statute?

You will find it frequently said in judicial opinions of my court and others that the judge's objective in interpreting a statute is to give effect to "the intent of the legislature." This principle, in one form or another, goes back at least as far as Blackstone. Unfortunately, it does not square with some of the (few) generally accepted concrete rules of statutory construction. One is the rule that when the text of a statute is clear, that is the end of the matter. Why should that be so, if what the legislature *intended*, rather than what it *said*, is the object of our inquiry? In selecting the words of the statute, the legislature might have misspoken. Why not permit that to be demonstrated from the floor debates? Or indeed, why not accept, as proper material for the court to consider, later explanations by the legislators—a sworn affidavit signed by the majority of each house, for example, as to what they *really* meant?

Another accepted rule of construction is that ambiguities in a newly enacted statute are to be resolved in such fashion as to make the statute, not only internally consistent, but also compatible with previously enacted laws. We simply assume, for purposes of our search for "intent," that the enacting legislature was aware of all those other laws. Well of course that is a fiction, and if we were really looking for the subjective intent of the enacting legislature we would more likely find it by paying attention to the text (and legislative history) of the new statute in isolation.

The evidence suggests that, despite frequent statements to the contrary, we do not really look for subjective legislative intent. We look for a sort of "objectified" intent—the intent that a reasonable person would gather from the text of the law, placed alongside the remainder of the *corpus juris*. . . . And the reason we adopt this objectified version is, I think, that it is simply incompatible with democratic government, or indeed, even with fair government, to have the meaning of a law determined by what the lawgiver meant, rather than by what the lawgiver promulgated. That seems to me one step worse than the trick the emperor Nero was said to engage in: posting edicts high up on the pillars, so that they could not easily be read. Government by unexpressed intent is similarly tyrannical. It is the *law* that governs, not the intent of the lawgiver. That seems to me the essence of the famous American ideal set forth in the Massachusetts constitution: A government of laws, not of men. Men may intend what they will; but it is only the laws that they enact which bind us.

In reality, however, if one accepts the principle that the object of judicial interpretation is to determine the intent of the legislature, being bound by genuine but unexpressed legislative intent rather

than the law is only the *theoretical* threat. The *practical* threat is that, under the guise or even the self-delusion of pursuing unexpressed legislative intents, common-law judges will in fact pursue their own objectives and desires, extending their lawmaking proclivities from the common law to the statutory field. When you are told to decide, not on the basis of what the legislature said, but on the basis of what it *meant*, and are assured that there is no necessary connection between the two, your best shot at figuring out what the legislature meant is to ask yourself what a wise and intelligent person *should* have meant; and that will surely bring you to the conclusion that the law means what you think it *ought* to mean—which is precisely how judges decide things under the common law.

Church of the Holy Trinity

To give some concrete form to the danger I warn against, let me describe what I consider to be the prototypical case involving the triumph of supposed "legislative intent" (a handy cover for judicial intent) over the text of the law. It is called *Church of the Holy Trinity* v. *United States* and was decided by the Supreme Court of the United States in 1892. The Church of the Holy Trinity, located in New York City, contracted with an Englishman to come over to be its rector and pastor. The United States claimed that this agreement violated a federal statute that made it unlawful for any person to "in any way assist or encourage the importation or migration of any alien . . . into the United States, . . . under contract or agreement . . . made previous to the importation or migration of such alien . . . , to perform labor or service of any kind in the United States. . . . " The Circuit Court for the Southern District of New York held the church liable for the fine that the statute provided. The Supreme Court reversed. . . .

The Court proceed[ed] to conclude from various extratextual indications, including even a snippet of legislative history (highly unusual in those days), that the statute was intended to apply only to *manual* labor—which renders the exceptions for actors, artists, lecturers, and singers utterly inexplicable. The Court then shifts gears and devotes the last seven pages of its opinion to a lengthy description of how and why we are a religious nation. That being so, it says, "[t]he construction invoked cannot be accepted as correct." . . .

Well of course I think that the act was within the letter of the statute, and was therefore within the statute: end of case. Congress can enact foolish statutes as well as wise ones, and it is not for the courts to decide which is which and rewrite the former. . . .

Another modern and forthright approach to according courts the power to revise statutes is set forth in Professor Eskridge's recent book, *Dynamic Statutory Interpretation.* The essence of it is acceptance of the proposition that it is proper for the judge who applies a statute to consider "not only what the statute means abstractly, or even on the basis of legislative history, but also what it ought to mean in terms of the needs and goals of our present day society." The law means what it ought to mean.

I agree . . . that many decisions can be cited which, by subterfuge, accomplish precisely what . . . Eskridge and other honest nontextualists propose. As I have said, "legislative intent" divorced from text is one of those subterfuges; and as I have described, *Church of the Holy Trinity* is one of those cases. What I think is needed, however, is not rationalization of this process but abandonment of it. It is simply not compatible with democratic theory that laws mean whatever they ought to mean, and that unelected judges decide what that is.

It may well be that the statutory interpretation adopted by the Court in *Church of the Holy Trinity* produced a desirable result; and it may even be (though I doubt it) that it produced the unexpressed result actually intended by Congress, rather than merely the one desired by the Court. Regardless, the decision was wrong because it failed to follow the text. The text is the law, and it is the text that must be observed. . . .

TEXTUALISM

The philosophy of interpretation I have described above is known as textualism. In some sophisticated circles, it is considered simpleminded—"wooden," "unimaginative," "pedestrian." It is none of that. To be a textualist in good standing, one need not be too dull to perceive the broader social purposes that

a statute is designed, or could be designed, to serve; or too hide-bound to realize that new times require new laws. One need only hold the belief that judges have no authority to pursue those broader purposes or write those new laws.

Textualism should not be confused with so-called strict constructionism, a degraded form of textualism that brings the whole philosophy into disrepute. I am not a strict constructionist, and no one ought to be—though better that, I suppose, than a nontextualist. A text should not be construed strictly, and it should not be construed leniently; it should be construed reasonably, to contain all that it fairly means. The difference between textualism and strict constructionism can be seen in a case my Court decided four terms ago. The statute at issue provided for an increased jail term if, "during and in relation to . . . [a] drug trafficking crime," the defendant "uses . . . a firearm." The defendant in this case had sought to purchase a quantity of cocaine; and what he had offered to give in exchange for the cocaine was an unloaded firearm, which he showed to the drug-seller. The Court held, I regret to say, that the defendant was subject to the increased penalty, because he had "used a firearm during and in relation to a drug trafficking crime." The vote was not even close (6–3). I dissented. Now I cannot say whether my colleagues in the majority voted the way they did because they are strict-construction textualists, or because they are not textualists at all. But a proper textualist, which is to say my kind of textualist, would surely have voted to acquit. The phrase "uses a gun" fairly connoted use of a gun for what guns are normally used for, that is, as a weapon. As I put the point in my dissent, when you ask someone, "Do you use a cane?" you are not inquiring whether he has hung his grandfather's antique cane as a decoration in the hallway.

But while the good textualist is not a literalist, neither is he a nihilist. Words do have a limited range of meaning, and no interpretation that goes beyond that range is permissible. My favorite example of a departure from text—and certainly the departure that has enabled judges to do more freewheeling lawmaking than any other—pertains to the Due Process Clause found in the Fifth and Fourteenth Amendments of the United States Constitution, which says that no person shall "be deprived of life, liberty, or property without due process of law." It has been interpreted to prevent the government from taking away certain liberties *beyond* those, such as freedom of speech and of religion, that are specifically named in the Constitution. (The first Supreme Court case to use the Due Process Clause in this fashion was, by the way, *Dred Scott*—not a desirable parentage.) Well, it may or may not be a good thing to guarantee additional liberties, but the Due Process Clause quite obviously does not bear that interpretation. By its inescapable terms, it guarantees only process. Property can be taken by the state; liberty can be taken; even life can be taken; but not without the *process* that our traditions require—notably, a validly enacted law and a fair trial. To say otherwise is to abandon textualism, and to render democratically adopted texts mere springboards for judicial lawmaking.

Of all the criticisms leveled against textualism, the most mindless is that it is "formalistic." The answer to that is, *of course it's formalistic!* The rule of law is *about* form. If, for example, a citizen performs an act—let us say the sale of certain technology to a foreign country—which is prohibited by a widely publicized bill proposed by the administration and passed by both houses of Congress, *but not yet signed by the President,* that sale is lawful. It is of no consequence that everyone knows both houses of Congress and the President wish to prevent that sale. Before the wish becomes a binding law, it must be embodied in a bill that passes both houses and is signed by the President. Is that not formalism? A murderer has been caught with blood on his hands, bending over the body of his victim; a neighbor with a video camera has filmed the crime; and the murderer has confessed in writing and on videotape. We nonetheless insist that before the state can punish this miscreant, it must conduct a full-dress criminal trial that results in a verdict of guilty. Is that not formalism? Long live formalism. It is what makes a government a government of laws and not of men. . . .

INTERPRETING CONSTITUTIONAL TEXTS

Without pretending to have exhausted the vast topic of textual interpretation, I wish to address a final subject: the distinctive problem of constitutional

interpretation. The problem is distinctive, not because special principles of interpretation apply, but because the usual principles are being applied to an unusual text. Chief Justice Marshall put the point as well as it can be put in *McCulloch* v. *Maryland:*

> A constitution, to contain an accurate detail of all the subdivisions of which its great powers will admit, and of all the means by which they may be carried into execution, would partake of the prolixity of a legal code, and could scarcely be embraced by the human mind. It would probably never be understood by the public. Its nature, therefore, requires, that only its great outlines should be marked, its important objects designated, and the minor ingredients which compose those objects be deduced from the nature of the objects themselves.

In textual interpretation, context is everything, and the context of the Constitution tells us not to expect nit-picking detail, and to give words and phrases an expansive rather than narrow interpretation—though not an interpretation that the language will not bear.

Take, for example, the provision of the First Amendment that forbids abridgment of "the freedom of speech, or of the press." That phrase does not list the full range of communicative expression. Handwritten letters, for example, are neither speech nor press. Yet surely there is no doubt they cannot be censored. In this constitutional context, speech and press, the two most common forms of communication, stand as a sort of synecdoche for the whole. That is not strict construction, but it is reasonable construction. . . .

[T]he Great Divide with regard to constitutional interpretation is not that between Framers' intent and objective meaning, but rather that between *original* meaning (whether derived from Framers' intent or not) and *current* meaning. The ascendant school of constitutional interpretation affirms the existence of what is called The Living Constitution, a body of law that (unlike normal statutes) grows and changes from age to age, in order to meet the needs of a changing society. And it is the judges who determine those needs and "find" that changing law. . . .

If you go into a constitutional law class, or study a constitutional law casebook, or read a brief filed in a constitutional law case, you will rarely find the discussion addressed to the text of the constitutional provision that is at issue, or to the question of what was the originally understood or even the originally intended meaning of that text. The starting point of the analysis will be Supreme Court cases, and the new issue will presumptively be decided according to the logic that those cases expressed, with no regard for how far that logic, thus extended, has distanced us from the original text and understanding. Worse still, however, it is known and understood that if that logic fails to produce what in the view of the current Supreme Court is the *desirable* result for the case at hand, then, like good common-law judges, the Court will distinguish its precedents, or narrow them, or if all else fails overrule them, in order that the Constitution might mean what it *ought* to mean. Should there be—to take one of the less controversial examples—a constitutional right to die? If so, there is. Should there be a constitutional right to reclaim a biological child put out for adoption by the other parent? Again, if so, there is. If it is good, it is so. Never mind the text that we are supposedly construing; we will smuggle these new rights in, if all else fails, under the Due Process Clause (which, as I have described, is textually incapable of containing them). Moreover, what the Constitution meant yesterday it does not necessarily mean today. As our opinions say in the context of our Eighth Amendment jurisprudence (the Cruel and Unusual Punishments Clause), its meaning changes to reflect "the evolving standards of decency that mark the progress of a maturing society." . . .

FLEXIBILITY AND LIBERALITY OF THE LIVING CONSTITUTION

The argument most frequently made in favor of The Living Constitution is a pragmatic one: Such an evolutionary approach is necessary in order to provide the "flexibility" that a changing society requires; the Constitution would have snapped if it had not been permitted to bend and grow. This might be a persuasive argument if most of the "growing" that the proponents of this approach have brought upon us in the past, and are determined to bring upon us in the future, were the

elimination of restrictions upon democratic government. But just the opposite is true. Historically, and particularly in the past thirty-five years, the "evolving" Constitution has imposed a vast array of new constraints—new inflexibilities—upon administrative, judicial, and legislative action. To mention only a few things that formerly could be done or not done, as the society desired, but now cannot be done:

- admitting in a state criminal trial evidence of guilt that was obtained by an unlawful search;
- permitting invocation of God at public-school graduations;
- electing one of the two houses of a state legislature the way the United States Senate is elected, i.e., on a basis that does not give all voters numerically equal representation;
- terminating welfare payments as soon as evidence of fraud is received, subject to restoration after hearing if the evidence is satisfactorily refuted;
- imposing property requirements as a condition of voting;
- prohibiting anonymous campaign literature;
- prohibiting pornography.

And the future agenda of constitutional evolutionists is mostly more of the same—the creation of *new* restrictions upon democratic government, rather than the elimination of old ones. *Less* flexibility in government, not *more.* As things now stand, the state and federal governments may either apply capital punishment or abolish it, permit suicide or forbid it—all as the changing times and the changing sentiments of society may demand. But when capital punishment is held to violate the Eighth Amendment, and suicide is held to be protected by the Fourteenth Amendment, all flexibility with regard to those matters will be gone. No, the reality of the matter is that, generally speaking, devotees of The Living Constitution do not seek to facilitate social change but to prevent it. . . .

Several terms ago a case came before the Supreme Court involving a prosecution for sexual abuse of a young child. The trial court found that the child would be too frightened to testify in the presence of the (presumed) abuser, and so, pursuant to state law, she was permitted to testify with only the prosecutor and defense counsel present, with the defendant, the judge, and the jury watching over closed-circuit television. A reasonable enough procedure, and it was held to be constitutional by my Court. I dissented, because the Sixth Amendment provides that "[i]n *all* criminal prosecutions the accused shall enjoy the right . . . to be confronted with the witnesses against him" (emphasis added). There is no doubt what confrontation meant—or indeed means today. It means face-to-face, not watching from another room. And there is no doubt what one of the major purposes of that provision was: to induce *precisely* that pressure upon the witness which the little girl found it difficult to endure. It is difficult to accuse someone to his face, particularly when you are lying. Now no extrinsic factors have changed since that provision was adopted in 1791. Sexual abuse existed then, as it does now; little children were more easily upset than adults, then as now; a means of placing the defendant out of sight of the witness existed then as now (a screen could easily have been erected that would enable the defendant to see the witness, but not the witness the defendant). But the Sixth Amendment nonetheless gave *all* criminal defendants the right to *confront* the witnesses against them, because that was thought to be an important protection. The only significant things that *have* changed, I think, are the society's sensitivity to so-called psychic trauma (which is what we are told the child witness in such a situation suffers) and the society's assessment of where the proper balance ought to be struck between the two extremes of a procedure that assures convicting 100 percent of all child abusers, and a procedure that assures acquitting 100 percent of those falsely accused of child abuse. I have no doubt that the society is, as a whole, happy and pleased with what my Court decided. But we should not pretend that the decision did not *eliminate* a liberty that previously existed.

LACK OF A GUIDING PRINCIPLE FOR EVOLUTION

. . . Perhaps the most glaring defect of Living Constitutionalism, next to its incompatibility with the whole anti-evolutionary purpose of a

constitution, is that there is no agreement, and no chance of agreement, upon what is to be the guiding principle of the evolution. . . . What is it that the judge must consult to determine when, and in what direction, evolution has occurred? Is it the will of the majority, discerned from newspapers, radio talk shows, public opinion polls, and chats at the country club? Is it the philosophy of Hume, or of John Rawls, or of John Stuart Mill, or of Aristotle? As soon as the discussion goes beyond the issue of whether the Constitution is static, the evolutionists divide into as many camps as there are individual views of the good, the true, and the beautiful. I think that is inevitably so, which means that evolutionism is simply not a practicable constitutional philosophy.

I do not suggest, mind you, that originalists always agree upon their answer. There is plenty of room for disagreement as to what original meaning was, and even more as to how that original meaning applies to the situation before the court. But the originalist at least knows what he is looking for: the original meaning of the text. Often—indeed, I dare say usually—that is easy to discern and simple to apply. Sometimes (though not very often) there will be disagreement regarding the original meaning; and sometimes there will be disagreement as to how that original meaning applies to new and unforeseen phenomena. How, for example, does the First Amendment guarantee of "the freedom of speech" apply to new technologies that did not exist when the guarantee was created—to sound trucks, or to government-licensed over-the-air television? In such new fields the Court must follow the trajectory of the First Amendment, so to speak, to determine what it requires—and assuredly that enterprise is not entirely cut-and-dried but requires the exercise of judgment.

But the difficulties and uncertainties of determining original meaning and applying it to modern circumstances are negligible compared with the difficulties and uncertainties of the philosophy which says that the Constitution *changes*; that the very act which it once prohibited it now permits, and which it once permitted it now forbids; and that the key to that change is unknown and unknowable. The originalist, if he does not have all the answers, has many of them. . . . For the evolutionist, on the other hand, every question is an open question, every day a new day. No fewer than three of the Justices with whom I have served have maintained that the death penalty is unconstitutional, *even though its use is explicitly contemplated in the Constitution.* The Due Process Clause of the Fifth and Fourteenth Amendments says that no person shall be deprived of life without due process of law; and the Grand Jury Clause of the Fifth Amendment says that no person shall be held to answer for a capital crime without grand jury indictment. No matter. Under The Living Constitution the death penalty may have *become* unconstitutional. And it is up to each Justice to decide for himself (under no standard I can discern) when that occurs.

In the last analysis, however, it probably does not matter what principle, among the innumerable possibilities, the evolutionist proposes to determine in what direction The Living Constitution will grow. Whatever he might propose, at the end of the day an evolving constitution will evolve the way the majority wishes. The people will be willing to leave interpretation of the Constitution to lawyers and law courts so long as the people believe that it is (like the interpretation of a statute) essentially lawyers' work—requiring a close examination of text, history of the text, traditional understanding of the text, judicial precedent, and so forth. But if the people come to believe that the Constitution is *not* a text like other texts; that it means, not what it says or what it was understood to mean, but what it *should* mean, in light of the "evolving standards of decency that mark the progress of a maturing society"—well, then, they will look for qualifications other than impartiality, judgment, and lawyerly acumen in those whom they select to interpret it. More specifically, they will look for judges who agree with *them* as to what the evolving standards have evolved to; who agree with *them* as to what the Constitution *ought* to be.

It seems to me that that is where we are heading, or perhaps even where we have arrived. Seventy-five years ago, we believed firmly enough in a rock-solid, unchanging Constitution that we felt it necessary to adopt the Nineteenth Amendment to give women the vote. The battle was not fought in the courts, and few thought that it could be, despite the constitutional guarantee of Equal Protection of the Laws; that provision did not, when it was adopted, and hence did not in 1920, guarantee equal access to the

ballot but permitted distinctions on the basis not only of age but of property and of sex. Who can doubt that if the issue had been deferred until today, the Constitution would be (formally) unamended, and the courts would be the chosen instrumentality of change? The American people have been converted to belief in The Living Constitution, a "morphing" document that means, from age to age, what it ought to mean. And with that conversion has inevitably come the new phenomenon of selecting and confirming federal judges, at all levels, on the basis of their views regarding a whole series of proposals for constitutional evolution. If the courts are free to write the Constitution anew, they will, by God, write it the way the majority wants; the appointment and confirmation process will see to that. This, of course, is the end of the Bill of Rights, whose meaning will be committed to the very body it was means to protect against: the majority. By trying to make the Constitution do everything that needs doing from age to age, we shall have caused it to do nothing at all.

REVIEW AND DISCUSSION QUESTIONS

1. What was decided in *Church of the Holy Trinity*? How would a textualist like Scalia have decided that case? How do you think the case should have been decided?
2. Although Scalia upholds "textualism," he rejects "strict constructionism." Explain.
3. Scalia argues that controversial constitutional clauses should be interpreted based on their original meaning. Assess that argument in light of Dworkin's distinction (see "'Vague' Constitutional Clauses") between a concept and a conception.
4. Scalia discusses the Sixth Amendment and the prosecution of defendants accused of child abuse. What did the Supreme Court decide, and why? Why did Scalia dissent? What should the Court have decided?
5. What point does Scalia make about the Nineteenth Amendment? Is he right that if the matter had been deferred until today, the courts would have settled it? Would it have been better or worse if the Supreme Court had granted women a right to vote without a constitutional amendment?
6. At the end of his essay, Scalia argues that adopting the ideal of a Living Constitution will lead to rewriting the Constitution in the way the majority wants. Is this consistent with his earlier argument that many recent Supreme Court decisions have wrongly limited what the majority can do? In your view, is the idea of a Living Constitution as dangerous as Scalia thinks, or is it necessary for sound and progressive legal decisions?

A Reponse to Scalia

Ronald Dworkin

What matters for Scalia, Dworkin argues, is semantic intention: what the legislators (or framers of the Constitution) intended to say in enacting the law (as opposed to what they intended or expected would be the consequence of their enacting it). Dworkin agrees with this. But when it comes to interpreting the abstract clauses of the Constitution, such as the Bill of Rights, this distinction points to two types of originalism. "Semantic" originalism insists those clauses should be read to say what those who made them intended to say. "Expectation" originalism argues that those clauses should be understood to have the consequences those who made them expected them to have. Dworkin argues that only the former sort of originalism makes sense or, indeed, squares with what the framers of the Constitution thought. But it requires us to make fresh judgments about how to apply the abstract principles of the Constitution to particular cases. By subordinating those judgments to a judgment about how the framers would have expected those principles to be interpreted, Scalia is unfaithful to his own theory of interpretation.

1

Justice Scalia . . . has set out with laudable clarity a sensible account of statutory interpretation . . . But I believe he has seriously misunderstood the implications of his general account for constitutional law, and that his lectures therefore have a schizophrenic character. He begins with a general theory that entails a style of constitutional adjudication which he ends by denouncing.

His initial argument rests on a crucial distinction between law and intention. "Men may intend what they will," he says, "but it is only the laws that they enact which bind us," and he is scornful of decisions like *Holy Trinity*, in which the Supreme Court, conceding that the "letter" of a statute forbade what the church had done, speculated that Congress did not intend that result. Indeed, he is skeptical about the very idea of a corporate legislative "intention"; most members of Congress, he says, have never thought about the unforeseen issues of interpretation that courts must face. A careless reader might object, however, that any coherent account of statutory interpretation *must* be based on assumptions about someone's (or somebody's) intention, and that Scalia's own account accepts this at several points. Scalia admits that courts should remedy "scrivener's error." He rejects "strict constructionism"—he thinks the Supreme Court's "literalist" decision in the "firearm" case, *Smith* v. *United States*, was silly. He credits at least some of the "canons" of interpretation as being an "indication" of meaning. And he says that it would be absurd to read the First Amendment's protection of speech and press as not applying to handwritten notes, which are, technically, neither.

Each of these clarifications allows respect for intention to trump literal text, and the careless objection I am imagining therefore claims an inconsistency. Scalia's defenders might say, in reply to the objection, that he is not an *extreme* textualist, and that these adjustments are only concessions to common sense and practicality. But that misunderstands the objection, which is that the concessions undermine Scalia's position altogether, because they recognize not only the intelligibility but the priority of legislative intention, both of which he begins by denying. If judges can appeal to a presumed legislative intent to add to the plain meaning of "speech" and "press," or to subtract from the plain meaning of "uses a firearm," why can they not appeal to the same legislative intent to allow a priest to enter the country? Scalia's answer to this objection must not rely on any self-destructive "practicality" claim. It must rely instead on a distinction between *kinds* of intention, a distinction he does not make explicitly, but that must lie at the heart of his theory if the theory is defensible at all.

This is the crucial distinction between what some officials intended to *say* in enacting the language they used, and what they intended—or expected or hoped—would be the *consequence* of their saying it. Suppose a boss tells his manager (without winking) to hire the most qualified applicant for a new job. The boss might think it obvious that his own son, who is an applicant, is the most qualified; indeed he might not have given the instruction unless he was confident that the manager would think so too. Nevertheless, what the boss *said*, and *intended* to say, was that the most qualified applicant should be hired, and if the manager thought some other applicant better qualified, but hired the boss's son to save his own job, he would not be following the standard the boss had intended to lay down.

So what I called the careless objection is wrong. The supposed lapses from Scalia's textualism it cites are not lapses at all, because textualism insists on deference to one kind of intention—semantic intention—and in all his remarks so far cited Scalia is deferring to that. Any reader of anything must attend to semantic intention, because the same sounds or even words can be used with the intention of saying different things. If I tell you (to use Scalia's own example) that I admire bays, you would have to decide whether I intended to say that I admire certain horses or certain bodies of water. Until you had, you would have no idea what I had actually said even though you would know what sounds I had uttered. The phrase "using a firearm" might naturally be used, in some contexts, with the intention of describing only situations in which a gun is used as a threat; the same phrase might be used, in other contexts, to mean using a gun for any purpose including barter. We do not know what Congress actually said, in using a

similar phrase, until we have answered the question of what it is reasonable to suppose, in all the circumstances including the rest of the statute, it intended to say in speaking as it did.

When we are trying to decide what someone meant to say, in circumstances like these, we are deciding which clarifying *translation* of his inscriptions is the best. It is a matter of complex and subtle philosophical argument what such translations consist in, and how they are possible—how, for example, we weave assumptions about what the speaker believes and wants, and about what it would be rational for him to believe and want, into decisions about what he meant to say. The difficulties are greatly increased when we are translating not the utterances of a real person but those of an institution like a legislature. We rely on personification—we suppose that the institution has semantic intentions of its own—and it is difficult to understand what sense that makes, or what special standards we should use to discover or construct such intentions. Scalia would not agree with my own opinions about these matters. But we do agree on the importance of the distinction I am emphasizing: between the question of what a legislature intended to say in the laws it enacted, which judges applying those laws must answer, and the question of what the various legislators as individuals expected or hoped the consequences of those laws would be, which is a very different matter.

Holy Trinity illustrates the difference and its importance. There can be no serious doubt that Congress meant to say what the words it used would naturally be understood to say. It is conceivable—perhaps even likely—that most members would have voted for an exception for English priests had the issue been raised. But that is a matter of (counterfactual) expectations, not of semantic intention. The law, as Scalia emphasizes, is what Congress has said, which is fixed by the best interpretation of the language it used, not by what some proportion of its members wanted or expected or assumed would happen, or would have wanted or expected or assumed if they had thought of the case. Not everyone agrees with that judgment. Some lawyers think that it accords better with democracy if judges defer to reasonable assumptions about what most legislators wanted or would have wanted, even when the language they used does not embody those actual or hypothetical wishes. After all, these lawyers argue, legislation should reflect what those who have been elected by the people actually think best for the country. Scalia disagrees with that judgment: he thinks it more democratic to give semantic intention priority over expectation intention when the two conflict, as they putatively did in *Holy Trinity*.

2

Now consider the implications of textualism so understood for the most important part of Scalia's judicial duties: interpreting the exceedingly abstract clauses of the Bill of Rights and later rights-bearing amendments. Scalia describes himself as a constitutional "originalist." But the distinction we made allows us a further distinction between two forms of originalism: "semantic" originalism, which insists that the rights-granting clauses be read to say what those who made them intended to say, and "expectation" originalism, which holds that these clauses should be understood to have the consequences that those who made them expected them to have. Consider, to see the difference, the *Brown* question: does the Fourteenth Amendment guarantee of "equal protection of the laws" forbid racial segregation in public schools? We know that the majority of the members of Congress who voted for that amendment did not expect or intend it to have that consequence: they themselves sustained racial segregation in the schools of the District of Columbia. So an expectation-originalist would interpret the Fourteenth Amendment to permit segregation and would declare the Court's decision wrong. But there is no plausible interpretation of what these statesmen meant to *say*, in laying down the language "equal protection of the laws," that entitles us to conclude that they *declared* segregation constitutional. On the contrary, as the Supreme Court held, the best understanding of their semantic intentions supposes that they meant to, and did, lay down a general principle of political morality which (it had become clear by 1954) condemns racial segregation. So, on that ground, a semantic-originalist would concur in the Court's decision.

If Scalia were faithful to his textualism, he would be a semantic-originalist. But is he? Notice

his brief discussion of whether capital punishment offends the Eighth Amendment's prohibition against "cruel and unusual" punishments. An expectation-originalist would certainly hold that it does not, for the reasons Scalia cites. The "framers" would hardly have bothered to stipulate that "life" may be taken only after due process if they thought that the Eighth Amendment made capital punishment unconstitutional anyway. But the question is far more complicated for a semantic-originalist. For he must choose between two clarifying translations—two different accounts of what the framers intended to *say* in the Eighth Amendment. The first reading supposes that the framers intended to say, by using the words "cruel and unusual," that punishments generally thought cruel at the time they spoke were to be prohibited—that is, that they would have expressed themselves more clearly if they had used the phrase "punishments widely regarded as cruel and unusual at the date of this enactment" in place of the misleading language they actually used. The second reading supposes that they intended to lay down an abstract principle forbidding whatever punishments are in fact cruel and unusual. Of course, if the correct translation is the first version, then capital punishment does not violate the Eighth Amendment. But if the second, principled, translation is a more accurate account of what they intended to say, the question remains open. Just as the manager in my story could only follow his boss's principled instruction by using his own judgment, so judges could then only apply the Eighth Amendment by deciding whether capital punishment is in fact cruel and has now become (as in fact it has become, at least among democracies) unusual.

The textual evidence Scalia cites would be irrelevant for a semantic-originalist who translated the Eighth Amendment in a principled rather than a concrete and dated way. There is no contradiction in the following set of claims. The framers of the Eighth Amendment laid down a principle forbidding whatever punishments are cruel and unusual. They did not themselves expect or intend that that principle would abolish the death penalty, so they provided that death could be inflicted only after due process. But it does not follow that the abstract principle they stated does not, contrary to their own expectation, forbid capital punishment. Suppose some legislature enacts a law forbidding the hunting of animals that are members of "endangered species" and then, later in its term, imposes special license requirements for hunting, among other animals, minks. We would assume that the members who voted for both provisions did not think that minks were endangered. But we would not be justified in concluding from that fact that, as a matter of law, minks were excluded from the ban even if they plainly *were* endangered. The latter inference would be an example of *Holy Trinity* thinking.

You will now understand my concern about Scalia's consistency. For he cites the view that capital punishment is unconstitutional as so obviously preposterous that it is cause for wonder that three justices who served with him actually held such an opinion. If he were an expectation-originalist, we would not be surprised at that view, or at the evidence he offers to support it. But for a semantic-originalist the question just *cannot* be foreclosed by references to the death penalty in the rest of the Constitution. A semantic-originalist would also have to think that the best interpretation of the Eighth Amendment was the dated rather than the principled translation, and even someone who might be drawn to that dated interpretation could not think the principled one *preposterous.*

On the contrary, it is the dated translation that seems bizarre. It is near inconceivable that sophisticated eighteenth-century statesmen, who were familiar with the transparency of ordinary moral language, would have used "cruel" as shorthand for "what we now think cruel." They knew how to be concrete when they intended to be: the various provisions for criminal and civil process in the Fourth, Fifth, Sixth, and Seventh Amendments do not speak of "fair" or "due" or "usual" procedures but lay down very concrete provisions. If they had intended a dated provision, they could and would have written an explicit one. Of course, we cannot imagine Madison or any of his contemporaries doing that: they wouldn't think it appropriate to protect what they took to be a fundamental right in such terms. But that surely means that the dated translation would be a plain mistranslation.

So Scalia's impatience with what seems the most natural statement of what the authors of the Eighth Amendment intended to say is puzzling. Part of the explanation may lie in his fear of what he calls a "morphing" theory of the Constitution—that the rights-bearing clauses are chameleons which change their meaning to conform to the needs and spirit of new times. He calls this chameleon theory "dominant," but it is hardly even intelligible, and I know of no prominent contemporary judge or scholar who holds anything like it. True, a metaphorical description of the Constitution as "living" has figured in constitutional rhetoric of the past, but this metaphor is much better understood as endorsing, not the chameleon theory, but the view I just described as the one that Scalia, if he were a semantic-originalist, might be expected to hold himself—that key constitutional provisions, as a matter of their original meaning, set out abstract principles rather than concrete or dated rules. If so, then the application of these abstract principles to particular cases, which takes fresh judgment, must be continually reviewed, not in an attempt to find substitutes for what the Constitution says, but out of respect for what it says.

I have defended that view in a series of books over the last decade, and some of what I have written might strike Scalia as saying that the Constitution itself changes, though I meant the opposite. I said, for example, that, subject to the constraints of integrity which require judges to keep faith with past decisions, "The Constitution insists that our judges do their best collectively to construct, reinspect, and revise, generation by generation, the skeleton of freedom and equality of concern that its great clauses, in their majestic abstraction, command."

It is that moral and principled reading of the Constitution, not the mythic chameleon claims he describes, that Scalia must produce reasons for rejecting. . . . Scalia argues that the First Amendment should be read not as abstract but as dated—that it should be read, that is, as guaranteeing only the rights it would have been generally understood to protect when it was enacted. . . .

[However,] Enlightenment statesmen were very unlikely to think that their own views represented the last word in moral progress. If they really were worried that future generations would protect rights less vigorously than they themselves did, they would have made plain that they intended to create a dated provision. . . . [Scalia] ignores, moreover, an apparently decisive argument against a translation of the First Amendment as dated. There *was* no generally accepted understanding of the right of free speech on which the framers could have based a dated clause even if they had wanted to write one. On the contrary, the disagreement about what that right comprises was much more profound when the amendment was enacted than it is now. When the dominant Federalist party enacted the Sedition Act in 1798, its members argued, relying on Blackstone, that "the freedom of speech" meant only freedom from "prior restraint"—in effect, freedom from an advance prohibition—and did not include any protection at all from punishment *after* publication. The opposing Republicans argued for a dramatically different view of the amendment. . . . All parties to the debate *assumed* that the First Amendment set out an abstract principle and that fresh judgment would be needed to interpret it. The Federalists relied, not on contemporary practice, which hardly supported their reading, but on the moral authority of Blackstone. The Republicans relied, not on contemporary practice either, but on the logic of freedom. No one supposed that the First Amendment codified some current and settled understanding, and the deep division among them showed that there was no settled understanding to codify.

So Scalia's discussion of the First Amendment is as puzzling as his briefer remarks about the Eighth Amendment. Now consider what he says about the Fourteenth Amendment's guarantee of "equal protection of the laws." He says that that clause "did not, when it was adopted, and hence did not in 1920, guarantee equal access to the ballot but permitted distinctions on the basis not only of age but of property and sex." Why is he so sure that the Equal Protection Clause did not always forbid discrimination on grounds of age, property, or sex (or, for that matter, sexual orientation)? Certainly when the amendment was adopted, few people *thought* that the clause had that consequence, any more than they thought that it had the consequence of making school segregation illegal. But the semantic-originalist would dismiss this as just what the framers and

later generations of lawyers expected, not a matter of what the framers actually *said*. If we look at the text they wrote, we see no distinction between racial discrimination and any other form of discrimination: the language is perfectly general, abstract, and principled. Scalia now reads into that language limitations that the language not only does not suggest but cannot bear, and he tries to justify this mistranslation by attributing understandings and expectations to statesmen that they may well have had, but that left no mark on the text they wrote. . . .

What has happened? Why does the resolute text-reader, dictionary-minder, expectation scorner of the beginning of these lectures change his mind when he comes to the most fundamental American statute of them all? He offers, in his final pages, an intriguing answer. He sees, correctly, that if we read the abstract clauses of the Bill of Rights as they were written—if we read them to say what their authors intended them to say rather than to deliver the consequences they expected them to have—then judges must treat these clauses as enacting abstract moral principles and must therefore exercise moral judgment in deciding what they *really* require. That does not mean ignoring precedent or textual or historical integrity or morphing the Constitution. It means, on the contrary, enforcing it in accordance with its text, in the only way that this can be done. Many conservative judges therefore reject semantic originalism as undemocratic; elected judges, they say, should not have that responsibility. Scalia gives nearly the opposite reason: he says the moral reading gives the people not too little but too much power, because it politicizes the appointment of Supreme Court justices and makes it more likely that justices will be appointed who reflect the changing moods of the majority. He fears that the constitutional rights of individuals will suffer.

History disagrees. Justices whose methods seem closest to the moral reading of the Constitution have been champions, not enemies, of individual rights, and, as the political defeat of Robert Bork's nomination taught us, the people seem content not only with the moral reading but with its individualist implications. . . .

It is, however, revealing that this is the scale on which Scalia finally wants his arguments to be weighed, and it may provide a final explanation, if not justification, for the inconsistency of his lectures as a whole. His most basic argument for textualism is drawn from majoritarian theory: he says that it is undemocratic when a statute is interpreted other than in accordance with the public text that was before legislators when they voted and is available to everyone in the community afterwards. His most basic argument for rejecting textualism in constitutional interpretation, on the other hand, reflects his *reservations* about majority rule. As with most of us, Scalia's attitudes about democracy are complex and ambivalent. I disagree with his judgment about which individual rights are genuine and important, and about whether the moral reading is a threat or an encouragement to freedom. But I agree with him that in the end the magnet of political morality is the strongest force in jurisprudence. The power of that magnet is nowhere more evident than in the rise and fall of his own love affair with textual fidelity.

REVIEW AND DISCUSSION QUESTIONS

1. What is the "careless objection" to Scalia, and how does the distinction that Dworkin draws between the two types of intention rescue Scalia from that objection?
2. What is the difference between "semantic" originalism and "expectation" originalism?
3. Dworkin contends that if Scalia were faithful to his textualism, he would be a semantic-originalist? Why? Do you agree or disagree?
4. Contrast the perspective of the semantic-originalist with that of the expectation-originalist with regard to the First Amendment, Eight Amendment, and Fourteenth Amendment. Do you agree or disagree with Dworkin about how those Amendments are to be interpreted.
5. Whose approach to constitutional interpretation is better—Scalia's or Dworkin's? Explain why.

The Second Amendment

District of Columbia v. Heller

Although it is only one sentence long, the Second Amendment is difficult to interpret, in part because its historical context seems so remote from the modern world. Moreover, the politically charged issue of gun control has often colored people's understanding of the Amendment. Unfortunately, the U.S. Supreme Court has rarely discussed the Second Amendment—in fact until *Heller* in 2008, it seems to have mentioned the Second Amendment in only three cases—in 1875, 1886, and 1939—and then only briefly.

In his majority opinion, Justice Scalia examines closely the language of the Second Amendment and how it would have been understood at the time. He concludes that the Second Amendment should be interpreted as guaranteeing an individual's right to possess a handgun in his or her home for purposes of self-defense and, thus, that the District of Columbia law banning the possession of handguns is unconstitutional. In his dissent, Justice Stevens argues that the Second Amendment protects the possession of firearms only in connection with service in a militia and, hence, does not rule out gun control measures like that of the District of Columbia. In a separate dissent, Justice Breyer argues that even if the Second Amendment does recognize an interest in possessing guns for purposes of self-defense unrelated to a militia, this interest has to be weighed against other legitimate interests, including the District of Columbia's interest in reducing handgun violence.

Justice Scalia delivered the opinion of the Court: We consider whether a District of Columbia prohibition on the possession of usable handguns in the home violates the Second Amendment to the Constitution.

I

The District of Columbia generally prohibits the possession of handguns. It is a crime to carry an unregistered firearm, and the registration of handguns is prohibited. . . . District of Columbia law also requires residents to keep their lawfully owned firearms, such as registered long guns, "unloaded and disassembled or bound by a trigger lock or similar device" unless they are located in a place of business or are being used for lawful recreational activities.

Respondent Dick Heller is a D. C. special police officer authorized to carry a handgun while on duty at the Federal Judicial Center. He applied for a registration certificate for a handgun that he wished to keep at home, but the District refused. He thereafter filed a lawsuit in the Federal District Court for the District of Columbia seeking, on Second Amendment grounds, to enjoin the city from enforcing the bar on the registration of handguns, the licensing requirement insofar as it prohibits the carrying of a firearm in the home without a license, and the trigger-lock requirement insofar as it prohibits the use of "functional firearms within the home." The District Court dismissed respondent's complaint. . . . The Court of Appeals for the District of Columbia Circuit, construing his complaint as seeking the right to render a firearm operable and carry it about his home in that condition only when necessary for self-defense, reversed. It held that the Second Amendment protects an individual right to possess firearms and that the city's total ban on handguns, as well as its requirement that firearms in the home be kept nonfunctional even when necessary for self-defense, violated that right.

II

We turn first to the meaning of the Second Amendment.

554 U.S.——(2008).

A

The Second Amendment provides: "A well regulated Militia, being necessary to the security of a free State, the right of the people to keep and bear Arms, shall not be infringed." . . .

The two sides in this case have set out very different interpretations of the Amendment. Petitioners and today's dissenting Justices believe that it protects only the right to possess and carry a firearm in connection with militia service. Respondent argues that it protects an individual right to possess a firearm unconnected with service in a militia, and to use that arm for traditionally lawful purposes, such as self-defense within the home.

The Second Amendment is naturally divided into two parts: its prefatory clause and its operative clause. The former does not limit the latter grammatically, but rather announces a purpose. The Amendment could be rephrased, "Because a well regulated Militia is necessary to the security of a free State, the right of the people to keep and bear Arms shall not be infringed." . . .

Logic demands that there be a link between the stated purpose and the command. The Second Amendment would be nonsensical if it read, "A well regulated Militia, being necessary to the security of a free State, the right of the people to petition for redress of grievances shall not be infringed." . . . But apart from that clarifying function, a prefatory clause does not limit or expand the scope of the operative clause. . . . Therefore, while we will begin our textual analysis with the operative clause, we will return to the prefatory clause to ensure that our reading of the operative clause is consistent with the announced purpose.

1. *Operative Clause.*

a. "Right of the People." The first salient feature of the operative clause is that it codifies a "right of the people." The unamended Constitution and the Bill of Rights use the phrase "right of the people" two other times, in the First Amendment's Assembly-and-Petition Clause and in the Fourth Amendment's Search-and-Seizure Clause. The Ninth Amendment uses very similar terminology . . . All three of these instances unambiguously refer to individual rights, not "collective" rights, or rights that may be exercised only through participation in some corporate body.

. . . [T]he "militia" in colonial America consisted of a subset of "the people"—those who were male, able bodied, and within a certain age range. Reading the Second Amendment as protecting only the right to "keep and bear Arms" in an organized militia therefore fits poorly with the operative clause's description of the holder of that right as "the people."

We start therefore with a strong presumption that the Second Amendment right is exercised individually and belongs to all Americans.

b. "Keep and bear Arms." We move now from the holder of the right—"the people"—to the substance of the right: "to keep and bear Arms."

Before addressing the verbs "keep" and "bear," we interpret their object: "Arms." The 18th-century meaning is no different from the meaning today. . . . The term was applied, then as now, to weapons that were not specifically designed for military use and were not employed in a military capacity. . . . Just as the First Amendment protects modern forms of communications, . . . and the Fourth Amendment applies to modern forms of search . . . the Second Amendment extends, prima facie, to all instruments that constitute bearable arms, even those that were not in existence at the time of the founding.

We turn to the phrases "keep arms" and "bear arms." . . . [T]he most natural reading of "keep Arms" in the Second Amendment is to "have weapons." . . . "Keep arms" was simply a common way of referring to possessing arms, for militiamen *and everyone else*. . . . At the time of the founding, as now, to "bear" meant to "carry." . . . From our review of founding-era sources, we conclude that this natural meaning was also the meaning that "bear arms" had in the 18th century. In numerous instances, "bear arms" was unambiguously used to refer to the carrying of weapons outside of an organized militia. . . .

c. Meaning of the Operative Clause. Putting all of these textual elements together, we find that they guarantee the individual right to possess and carry weapons in case of confrontation. This meaning is strongly confirmed by the historical background of the Second Amendment. We look to this because it has always been widely understood that the Second Amendment, like the First and Fourth Amendments, codified a *pre-existing* right. The very

text of the Second Amendment implicitly recognizes the pre-existence of the right and declares only that it "shall not be infringed." As we said in *United States* v. *Cruikshank*, 92 U. S. 542, 553 (1876), "[t]his is not a right granted by the Constitution. Neither is it in any manner dependent upon that instrument for its existence. The Second Amendment declares that it shall not be infringed. . . ."

Between the Restoration and the Glorious Revolution, the Stuart Kings Charles II and James II succeeded in using select militias loyal to them to suppress political dissidents, in part by disarming their opponents. . . . And, of course, what the Stuarts had tried to do to their political enemies, George III had tried to do to the colonists. In the tumultuous decades of the 1760s and 1770s, the Crown began to disarm the inhabitants of the most rebellious areas. That provoked polemical reactions by Americans invoking their rights as Englishmen to keep arms.

There seems to us no doubt, on the basis of both text and history, that the Second Amendment conferred an individual right to keep and bear arms. Of course the right was not unlimited, just as the First Amendment's right of free speech was not. . . . Thus, we do not read the Second Amendment to protect the right of citizens to carry arms for *any sort* of confrontation, just as we do not read the First Amendment to protect the right of citizens to speak for *any purpose.* Before turning to limitations upon the individual right, however, we must determine whether the prefatory clause of the Second Amendment comports with our interpretation of the operative clause.

2. *Prefatory Clause.* The prefatory clause reads: "A well regulated Militia, being necessary to the security of a free State. . . ."

a. "Well-Regulated Militia." In *United States* v. *Miller*, 307 U.S. 174, 179 (1939), we explained that "the Militia comprised all males physically capable of acting in concert for the common defense." That definition comports with founding-era sources. . . . [T]he adjective "well-regulated" implies nothing more than the imposition of proper discipline and training. . . .

b. "Security of a Free State." The phrase "security of a free State" meant "security of a free polity," not security of each of the several States. . . . It is true that the term "State" elsewhere in the Constitution refers to individual States, but the phrase "security of a free State" and close variations seem to have been terms of art in 18th-century political discourse, meaning a "free country" or free polity. . . .

There are many reasons why the militia was thought to be "necessary to the security of a free State." . . . First, of course, it is useful in repelling invasions and suppressing insurrections. Second, it renders large standing armies unnecessary . . . Third, when the able-bodied men of a nation are trained in arms and organized, they are better able to resist tyranny.

3. *Relationship between Prefatory Clause and Operative Clause.* We reach the question, then: Does the preface fit with an operative clause that creates an individual right to keep and bear arms? It fits perfectly, once one knows the history that the founding generation knew and that we have described above. That history showed that the way tyrants had eliminated a militia consisting of all the able-bodied men was not by banning the militia but simply by taking away the people's arms, enabling a select militia or standing army to suppress political opponents. This is what had occurred in England that prompted codification of the right to have arms in the English Bill of Rights.

The debate with respect to the right to keep and bear arms, as with other guarantees in the Bill of Rights, was not over whether it was desirable (all agreed that it was) but over whether it needed to be codified in the Constitution. During the 1788 ratification debates, the fear that the federal government would disarm the people in order to impose rule through a standing army or select militia was pervasive in Antifederalist rhetoric. . . .

It is therefore entirely sensible that the Second Amendment's prefatory clause announces the purpose for which the right was codified: to prevent elimination of the militia. The prefatory clause does not suggest that preserving the militia was the only reason Americans valued the ancient right; most undoubtedly thought it even more important for self-defense and hunting. But the threat that the new Federal Government would destroy the citizens' militia by taking away their arms was the reason that right—unlike some other English rights—was codified in a written Constitution. . . .

B

We now ask whether any of our precedents forecloses the conclusions we have reached about the meaning of the Second Amendment. . . .

JUSTICE STEVENS places overwhelming reliance upon this Court's decision in *United States* v. *Miller*, 307 U.S. 174 (1939) . . . [t]hat the Second Amendment "protects the right to keep and bear arms for certain military purposes, but that it does not curtail the legislature's power to regulate the nonmilitary use and ownership of weapons."

Nothing so clearly demonstrates the weakness of JUSTICE STEVENS' case. *Miller* did not hold that and cannot possibly be read to have held that. The judgment in the case upheld against a Second Amendment challenge two men's federal convictions for transporting an unregistered short-barreled shotgun in interstate commerce, in violation of the National Firearms Act, 48 Stat. 1236. It is entirely clear that the Court's basis for saying that the Second Amendment did not apply was *not* that the defendants were "bear[ing] arms" not "for . . . military purposes" but for "nonmilitary use." Rather, it was that the *type of weapon at issue* was not eligible for Second Amendment protection: "In the absence of any evidence tending to show that the possession or use of a [short-barreled shotgun] at this time has some reasonable relationship to the preservation or efficiency of a well regulated militia, we cannot say that the Second Amendment guarantees the right to keep and bear *such an instrument.*" "Certainly," the Court continued, "it is not within judicial notice that this weapon is any part of the ordinary military equipment or that its use could contribute to the common defense." Beyond that, the opinion provided no explanation of the content of the right.

This holding is not only consistent with, but positively suggests, that the Second Amendment confers an individual right to keep and bear arms (though only arms that "have some reasonable relationship to the preservation or efficiency of a well regulated militia"). Had the Court believed that the Second Amendment protects only those serving in the militia, it would have been odd to examine the character of the weapon rather than simply note that the two crooks were not militiamen. . . .

III

Like most rights, the right secured by the Second Amendment is not unlimited. . . . For example, the majority of the 19th-century courts to consider the question held that prohibitions on carrying concealed weapons were lawful under the Second Amendment or state analogues. . . . [N]othing in our opinion should be taken to cast doubt on longstanding prohibitions on the possession of firearms by felons and the mentally ill, or laws forbidding the carrying of firearms in sensitive places such as schools and government buildings, or laws imposing conditions and qualifications on the commercial sale of arms.

We also recognize another important limitation on the right to keep and carry arms. *Miller* said, as we have explained, that the sorts of weapons protected were those "in common use at the time." We think that limitation is fairly supported by the historical tradition of prohibiting the carrying of "dangerous and unusual weapons."

It may be objected that if weapons that are most useful in military service—M-16 rifles and the like—may be banned, then the Second Amendment right is completely detached from the prefatory clause. But . . . the conception of the militia at the time of the Second Amendment's ratification was the body of all citizens capable of military service, who would bring the sorts of lawful weapons that they possessed at home to militia duty. It may well be true today that a militia, to be as effective as militias in the 18th century, would require sophisticated arms that are highly unusual in society at large. Indeed, it may be true that no amount of small arms could be useful against modern-day bombers and tanks. But the fact that modern developments have limited the degree of fit between the prefatory clause and the protected right cannot change our interpretation of the right.

IV

We turn finally to the law at issue here. As we have said, the law totally bans handgun possession in the home. It also requires that any lawful firearm in the home be disassembled or bound by a trigger lock at all times, rendering it inoperable. . . .

[W]e hold that the District's ban on handgun possession in the home violates the Second Amendment, as does its prohibition against rendering any lawful firearm in the home operable for the purpose of immediate self-defense. Assuming that Heller is not disqualified from the exercise of Second Amendment rights, the District must permit him to register his handgun and must issue him a license to carry it in the home.

***

We are aware of the problem of handgun violence in this country, and we take seriously the concerns raised by the many *amici* who believe that prohibition of handgun ownership is a solution. The Constitution leaves the District of Columbia a variety of tools for combating that problem, including some measures regulating handguns. But the enshrinement of constitutional rights necessarily takes certain policy choices off the table. These include the absolute prohibition of handguns held and used for self-defense in the home. Undoubtedly some think that the Second Amendment is outmoded in a society where our standing army is the pride of our Nation, where well-trained police forces provide personal security, and where gun violence is a serious problem. That is perhaps debatable, but what is not debatable is that it is not the role of this Court to pronounce the Second Amendment extinct.

We affirm the judgment of the Court of Appeals.

It is so ordered.

JUSTICE STEVENS, with whom JUSTICE SOUTER, JUSTICE GINSBURG, and JUSTICE BREYER join, dissenting: . . . The Second Amendment was adopted to protect the right of the people of each of the several States to maintain a well-regulated militia. It was a response to concerns raised during the ratification of the Constitution that the power of Congress to disarm the state militias and create a national standing army posed an intolerable threat to the sovereignty of the several States. Neither the text of the Amendment nor the arguments advanced by its proponents evidenced the slightest interest in limiting any legislature's authority to regulate private civilian uses of firearms. Specifically, there is no indication that the Framers of the Amendment intended to enshrine the common-law right of self-defense in the Constitution.

In 1934, Congress enacted the National Firearms Act, the first major federal firearms law. Upholding a conviction under that Act, this Court held that, "[i]n the absence of any evidence tending to show that possession or use of a 'shotgun having a barrel of less than eighteen inches in length' at this time has some reasonable relationship to the preservation or efficiency of a well regulated militia, we cannot say that the Second Amendment guarantees the right to keep and bear such an instrument." *Miller*, 307 U.S., at 178. The view of the Amendment we took in *Miller*—that it protects the right to keep and bear arms for certain military purposes, but that it does not curtail the Legislature's power to regulate the nonmilitary use and ownership of weapons—is both the most natural reading of the Amendment's text and the interpretation most faithful to the history of its adoption.

Since our decision in *Miller*, hundreds of judges have relied on the view of the Amendment we endorsed there; we ourselves affirmed it in 1980. See *Lewis* v. *United States*, 445 U.S. 55, 65–66, n. 8 (1980). No new evidence has surfaced since 1980 supporting the view that the Amendment was intended to curtail the power of Congress to regulate civilian use or misuse of weapons. Indeed, a review of the drafting history of the Amendment demonstrates that its Framers *rejected* proposals that would have broadened its coverage to include such uses.

The opinion the Court announces today fails to identify any new evidence supporting the view that the Amendment was intended to limit the power of Congress to regulate civilian uses of weapons. . . .

Even if the textual and historical arguments on both sides of the issue were evenly balanced, respect for the well-settled views of all of our predecessors on this Court, and for the rule of law itself . . . would prevent most jurists from endorsing such a dramatic upheaval in the law.

I

The text of the Second Amendment is brief. It provides:

"A well regulated Militia, being necessary to the security of a free State, the right of the people to keep and bear Arms, shall not be infringed."

Three portions of that text merit special focus: the introductory language defining the Amendment's purpose, the class of persons encompassed within its reach, and the unitary nature of the right that it protects.

"A well regulated Militia, being necessary to the security of a free State"

The preamble to the Second Amendment makes three important points. It identifies the preservation of the militia as the Amendment's purpose; it explains that the militia is necessary to the security of a free State; and it recognizes that the militia must be "well regulated." In all three respects it is comparable to provisions in several State Declarations of Rights that were adopted roughly contemporaneously with the Declaration of Independence. Those state provisions highlight the importance members of the founding generation attached to the maintenance of state militias; they also underscore the profound fear shared by many in that era of the dangers posed by standing armies. While the need for state militias has not been a matter of significant public interest for almost two centuries, that fact should not obscure the contemporary concerns that animated the Framers. . . .

"The right of the people"

The centerpiece of the Court's textual argument is its insistence that the words "the people" as used in the Second Amendment must have the same meaning, and protect the same class of individuals, as when they are used in the First and Fourth Amendments. According to the Court, in all three provisions—as well as the Constitution's preamble, section 2 of Article I, and the Tenth Amendment—"the term unambiguously refers to all members of the political community, not an unspecified subset." But the Court *itself* reads the Second Amendment to protect a "subset" significantly narrower than the class of persons protected by the First and Fourth Amendments; when it finally drills down on the substantive meaning of the Second Amendment, the Court limits the protected class to "law-abiding, responsible citizens." But the class of persons protected by the First and Fourth Amendments is *not* so limited; for even felons (and presumably irresponsible citizens as well) may invoke the protections of those constitutional provisions. The Court offers no way to harmonize its conflicting pronouncements. . . .

[T]he words "the people" in the Second Amendment refer back to the object announced in the Amendment's preamble. They remind us that it is the collective action of individuals having a duty to serve in the militia that the text directly protects and, perhaps more importantly, that the ultimate purpose of the Amendment was to protect the States' share of the divided sovereignty created by the Constitution. . . .

"To keep and bear Arms"

Although the Court's discussion of these words treats them as two "phrases"—as if they read "to keep" and "to bear"—they describe a unitary right: to possess arms if needed for military purposes and to use them in conjunction with military activities. . . .

[T]he clause protects only one right, rather than two. It does not describe a right "to keep arms" and a separate right "to bear arms." Rather, the single right that it does describe is both a duty and a right to have arms available and ready for military service, and to use them for military purposes when necessary. Different language surely would have been used to protect nonmilitary use and possession of weapons from regulation if such an intent had played any role in the drafting of the Amendment.

II

Until today, it has been understood that legislatures may regulate the civilian use and misuse of firearms so long as they do not interfere with the preservation of a well-regulated militia. The Court's announcement of a new constitutional right to own and use firearms for private purposes upsets that settled understanding, but leaves for future cases the formidable task of defining the scope of permissible regulations. . . .

For these reasons, I respectfully dissent.

JUSTICE BREYER, with whom JUSTICE STEVENS, JUSTICE SOUTER, and JUSTICE GINSBURG join, dissenting: We must decide whether a District of Columbia law that prohibits the possession of handguns in the home violates the Second Amendment. The majority, relying upon its view that the Second Amendment seeks to protect a right of personal self-defense, holds that this law violates that Amendment. In my view, it does not.

I

The majority's conclusion is wrong for two independent reasons. The first reason is that set forth by JUSTICE STEVENS—namely, that the Second Amendment protects militia-related, not self-defense-related, interests. These two interests are sometimes intertwined. To assure 18th-century citizens that they could keep arms for militia purposes would necessarily have allowed them to keep arms that they could have used for self-defense as well. But self-defense alone, detached from any militia-related objective, is not the Amendment's concern.

The second independent reason is that the protection the Amendment provides is not absolute. The Amendment permits government to regulate the interests that it serves. Thus, irrespective of what those interests are—whether they do or do not include an independent interest in self-defense—the majority's view cannot be correct unless it can show that the District's regulation is unreasonable or inappropriate in Second Amendment terms. This the majority cannot do.

In respect to the first independent reason, I agree with JUSTICE STEVENS, and I join his opinion. In this opinion I shall focus upon the second reason. . . . [T]he District's law is consistent with the Second Amendment even if that Amendment is interpreted as protecting a wholly separate interest in individual self-defense. That is so because the District's regulation, which focuses upon the presence of handguns in high-crime urban areas, represents a permissible legislative response to a serious, indeed life-threatening, problem.

Thus I here assume that one objective (but, as the majority concedes . . . not the *primary* objective) of those who wrote the Second Amendment was to help assure citizens that they would have arms available for purposes of self-defense. Even so, a legislature could reasonably conclude that the law will advance goals of great public importance, namely, saving lives, preventing injury, and reducing crime. The law is tailored to the urban crime problem in that it is local in scope and thus affects only a geographic area both limited in size and entirely urban; the law concerns handguns, which are specially linked to urban gun deaths and injuries, and which are the overwhelmingly favorite weapon of armed criminals; and at the same time, the law imposes a burden upon gun owners that seems proportionately no greater than restrictions in existence at the time the Second Amendment was adopted. In these circumstances, the District's law falls within the zone that the Second Amendment leaves open to regulation by legislatures.

II

. . . [T]he majority ignores an important question: Given the purposes for which the Framers enacted the Second Amendment, how should it be applied to modern-day circumstances that they could not have anticipated? Assume, for argument's sake, that the Framers did intend the Amendment to offer a degree of self-defense protection. Does that mean that the Framers also intended to guarantee a right to possess a loaded gun near swimming pools, parks, and playgrounds? That they would not have cared about the children who might pick up a loaded gun on their parents' bedside table? That they . . . would have lacked concern for the risk of accidental deaths or suicides that readily accessible loaded handguns in urban areas might bring? Unless we believe that they intended future generations to ignore such matters, answering questions such as the questions in this case requires judgment—judicial judgment exercised within a framework for constitutional analysis that guides that judgment and which makes its exercise transparent. One cannot answer those questions by combining inconclusive historical research with judicial *ipse dixit.*

The argument about method, however, is by far the less important argument surrounding today's decision. Far more important are the unfortunate

consequences that today's decision is likely to spawn. Not least of these, as I have said, is the fact that the decision threatens to throw into doubt the constitutionality of gun laws throughout the United States. I can find no sound legal basis for launching the courts on so formidable and potentially dangerous a mission. In my view, there simply is no untouchable constitutional right guaranteed by the Second Amendment to keep loaded handguns in the house in crime-ridden urban areas.

III

For these reasons, I conclude that the District's measure is a proportionate, not a disproportionate, response to the compelling concerns that led the District to adopt it. And, for these reasons as well as the independently sufficient reasons set forth by JUSTICE STEVENS, I would find the District's measure consistent with the Second Amendment's demands.

With respect, I dissent.

REVIEW AND DISCUSSION QUESTIONS

1. What are the facts of this case?
2. As Justice Scalia sees it, what is the relationship between the "prefatory" clause and the "operative" clause of the Second Amendment? What was the purpose of the Amendment? In your view, does that purpose restrict the right to keep and bear arms?
3. Contrast Justice Scalia's and Justice Stevens's interpretations of the phrases "right of the people" and "keep and bear arms."
4. Why did the framers believe that a militia was necessary for the security of a free state?
5. What did the Supreme Court decide in *United States v. Miller* (1939)? Does that decision support or cut against the majority opinion in Heller?
6. According to the majority, is the right to keep and bear arms unlimited?
7. In his dissent, Justice Breyer contends that the majority opinion is wrong for two reasons. What are they? How does his argument differ from that of Justice Stevens?

In Defense of Looseness: The Supreme Court and Gun Control

Richard A Posner

In this critique of *Heller*, federal appellate judge and legal theorist Richard A. Posner argues that Scalia's "originalist" method should have yielded the opposite result. Moreover, true originalism licenses loose (or flexible or nonliteral) construction because this was the reigning theory of legislative interpretation at the time. The older the constitutional provision and the more the environment has changed, the more appropriate is the method of loose construction. However, it is one thing to use loose construction to prevent the Constitution from being a straightjacket and another thing to use it to make the Constitution a straightjacket. Like *Roe v. Wade, Heller* imposes a uniform rule nationwide, giving short shrift to federalism and related values. Moreover it will lead to increased litigation over gun control. Judge Posner concludes by arguing that ideology, rather than principle, seems to drive Supreme Court decisions when liberal and conservative values collide.

At the end of June, the Supreme Court, in a case called *District of Columbia v. Heller*, invalidated the District's ban on the private ownership of pistols. It did so in the name of the Second Amendment to the Constitution. The decision was the most noteworthy of the Court's recent term. It is questionable in both method and result, and it is evidence that the Supreme Court, in deciding constitutional cases, exercises a freewheeling discretion strongly flavored with ideology.

The majority opinion, by Justice Antonin Scalia, concluded that the original, and therefore the authoritative, meaning of the Second Amendment is that Americans are entitled to possess pistols (and perhaps other weapons) for the defense of their homes. Scalia's entire analysis rests on this interpretive method, which denies the legitimacy of flexible interpretation designed to adapt the Constitution (so far as the text permits) to current conditions. The irony is that the "originalist" method would have yielded the opposite result.

The Second Amendment, part of the Bill of Rights added to the original Constitution in 1791, states: "A well regulated Militia being necessary to the security of a free State, the right of the people to keep and bear Arms, shall not be infringed." In other words: since a militia, provided that it is well regulated, is a very good thing for a free state to have, the federal government must not be allowed to castrate it by forbidding the people of the United States to possess weapons. For then the militia would have no weapons, and an unarmed militia is an oxymoron.

Politically conscious Americans in the late eighteenth century feared standing armies, having fought the British army in the Revolution, and feared centralized government (as in Britain); and on both counts they wanted to make sure that the states would be allowed to have armed militias. The federal government could regulate them but not disarm them. The fear was that in the absence of such a provision in the Bill of Rights, the provision in Article I of the Constitution authorizing Congress to organize, arm, discipline, and call into service "the Militia" (a term that embraces the state militias, because the same provision reserves the right to train and officer "the Militia" to the respective states) would enable Congress to disarm them. That fear surfaced in the debates over the ratification of the original Constitution and was, as Justice John Paul Stevens's dissenting opinion explains, the motivation for the Second Amendment.

The text of the amendment, whether viewed alone or in light of the concerns that actuated its adoption, creates no right to the private possession of guns for hunting or other sport, or for the defense of person or property. It is doubtful that the amendment could even be thought to require that members of state militias be allowed to keep weapons in their homes, since that would reduce the militias' effectiveness. Suppose part of a state's militia was engaged in combat and needed additional weaponry. Would the militia's commander have to collect the weapons from the homes of militiamen who had not been mobilized, as opposed to obtaining them from a storage facility? Since the purpose of the Second Amendment, judging from its language and background, was to assure the effectiveness of state militias, an interpretation that undermined their effectiveness by preventing states from making efficient arrangements for the storage and distribution of military weapons would not make sense.

The Court evaded the issue in *Heller* by cutting loose the Second Amendment from any concern with state militias (the "National Guard," as they are now called). The majority opinion acknowledges that allowing people to keep guns in their homes cannot help the militias, because modern military weapons are not appropriate for home defense (most of them are too dangerous), and anyway the opinion says that the only weapons the Second Amendment entitles people to possess are ones that are not "highly unusual in society at large." Modern military weapons are highly unusual in society at large. By creating a privilege to own guns of no interest to a militia, the Court decoupled the amendment's two clauses.

It justified this decoupling by arguing that the word "people" in the expression "the right of the people to keep and bear Arms" (the amendment's second clause) must encompass more than just militiamen, because eighteenth-century militias enrolled only able-bodied free men—a mere subset of the people of the United States. But obviously the Framers did not mean to confer even a prima facie constitutional right to possess guns on slaves, criminals, lunatics, and children. The purpose of the first clause of the amendment, the militia clause, is to narrow the right that the second clause confers on the "people."

I

My analysis to this point has been "originalist"—and it has led to the opposite conclusion from that of the majority of the Supreme Court. It has been a

narrow originalism, like that of Scalia's majority opinion, because it has ignored the interpretive conventions of the legal culture in which the Second Amendment was drafted and ratified. The reigning theory of legislative interpretation in the eighteenth century was loose (or flexible, or nonliteral) construction. This is explicit in William Blackstone's *Commentaries on the Laws of England*, on which the majority opinion in *Heller* ironically relies. In the *Commentaries* we read that a medieval law of Bologna stating that "whoever drew blood in the streets should be punished with the utmost severity" should not be interpreted to make punishable a surgeon "who opened the vein of a person that fell down in the street with a fit." Blackstone explained that "the fairest and most rational method to interpret the will of the legislator, is by exploring his intentions at the time when the law was made, by signs the most natural and probable. And these signs are either the words, the context, the subject matter, the effects and consequence, or the *spirit and reason of the law*. . . . As to the effects and consequence, the rule is, where words bear either none, or a very absurd signification, if literally understood, *we must a little deviate from the received sense of them*" (emphasis added). John Marshall, the greatest Supreme Court justice of the generation that wrote the Constitution and the Bill of Rights, was also a loose constructionist.

Originalism without the interpretive theory that the Framers and the ratifiers of the Constitution expected the courts to use in construing constitutional provisions is faux originalism. True originalism licenses loose construction. And loose construction is especially appropriate for interpreting a constitutional provision ratified more than two centuries ago, dealing with a subject that has been transformed in the intervening period by social and technological change, including urbanization and a revolution in warfare and weaponry.

The Framers of the Bill of Rights could not have been thinking of the crime problem in the large crime-ridden metropolises of twenty-first-century America, and it is unlikely that they intended to freeze American government two centuries hence at their eighteenth-century level of understanding. Because of the difficulty of amending the Constitution, it has from the beginning been loosely construed so as not to become a straitjacket or a suicide pact. The older the constitutional provision and the more the environment has changed since enactment, the more appropriate is the method of loose construction.

There are few more antiquated constitutional provisions than the Second Amendment. For example, the Framers and the ratifiers of the amendment probably did think that the right of militiamen to keep and bear arms entitled them to keep their weapons in their homes. They were expected to provide the militia's muskets rather than receive them from a weapons depot. Back then, moreover, guns required a lot of upkeep. Without constant care, they would rust and the powder would become moist. Storing the guns at a warehouse would have left many of them inoperable. To use this "original" understanding to allow members of the National Guard to store military weapons (machine guns, grenades, Hummers, and so on) would be preposterous, and it is disclaimed in the majority opinion.

II

In these and other ways, the *Heller* decision is exposed as an example of loose construction—despite the Court's pretense of engaging in originalist interpretation (but again, an originalism stripped of the original understanding of how a constitutional provision should be interpreted). . . .

The true springs of the *Heller* decision must be sought elsewhere than in the majority's declared commitment to originalism. The idea behind the decision—it is not articulated, of course, and perhaps not even consciously held—may simply be that turnabout is fair play. Liberal judges have used loose construction to expand constitutional prohibitions beyond any reasonable construal of original meaning; and now it is the conservatives' turn. Another plausible example of payback is the conservative justices' expansive interpretation of the free-speech clause of the First Amendment to limit regulation of campaign financing.

It is possible that in both the gun-control case and the campaign-finance cases the justices in the majority, rather than playing tit for tat, thought the laws they were invalidating very dumb, and in the case of the District of Columbia's ban on possession of

pistols thought the law wimpish and paternalistic, like requiring bikers to wear helmets. A law that bans possession of pistols outright may even be inferior, at least as a method of controlling crime, to a law that combines strict permit requirements with heavy penalties for violating them, or even to one that simply imposes draconian penalties on crimes committed with guns. But judges are not supposed to invalidate laws merely because, as legislators, they would have voted against them.

There is an important difference, obvious but often overlooked, between using loose construction to prevent making the Constitution a straitjacket and using it to make the Constitution a straitjacket. In *Kennedy v. Louisiana*, a decision handed down shortly before *Heller*, the Supreme Court held that to execute a person who rapes a child but does not kill her violates the cruel andunusual punishments clause of the Eighth Amendment. That was a loose construction that tied the hands of the states and the federal government, and Scalia and the other conservative justices dissented. But in *Heller* it was the liberal justices who were dissenting from a decision that ties the hands of the federal government, and of the states, too, if the Supreme Court decides that the Second Amendment constrains state as well as federal government action. Compare these two cases to the *Zelman* case, decided several years ago. There the Court upheld, against a challenge based on the clause of the First Amendment that forbids governmental establishments of religion, the funneling of public monies to private schools by means of vouchers that parents can use to pay for their kids' tuition. Most private schools are Catholic parochial schools. The interpretation of the establishment clause that permitted the use of public moneys to finance parochial schools rejected the imposition on government of a constitutional restraint that the liberal justices wanted to impose.

Another illuminating contrast to *Heller* is the recent *Kelo* decision. The Supreme Court held that the just compensation clause of the Fifth Amendment does not forbid a state to condemn private property and, having thus seized it, to turn it over to a private developer. The decision provoked outrage by conservatives, who oppose condemnation because it infringes rights of private property. They should not have been outraged. All the Court did was unshackle government from a potential constitutional constraint, and by doing so toss the issue into the political arena. And sure enough, in the wake of the decision a number of states, under pressure from property interests, curtailed their eminent domain powers.

Similarly, had the Supreme Court upheld the District of Columbia gun ordinance, it would not have been outlawing the private possession of guns. It would merely have been leaving the issue of gun control to the political process. The popularity of the decision and its prompt endorsement by both presidential candidates attests to the political power of the "gun lobby"; and an unpopular decision in favor of the government would actually have strengthened the lobby, just as *Roe v. Wade* strengthened the antiabortion movement. The proper time for using loose construction to enlarge constitutional restrictions on government action is when the group seeking the enlargement does not have good access to the political process to protect its interests, as abortion advocates, like gun advocates, did and do.

III

Constitutional interpretations that relax rather than tighten the Constitution's grip on the legislative and executive branches of government are especially welcome when there are regional or local differences in relevant conditions or in public opinion. The failure to recognize this point (or perhaps indifference to it) was the mistake that the Supreme Court made when it nationalized abortion rights in *Roe v. Wade.* It would be the mistake the Court would be making in the unlikely event that it created a federal constitutional right of homosexual marriage. It is the mistake the Court has made in *Heller.* The differences in attitudes toward private ownership of pistols across regions of the country and, outside the South, between urban and rural areas, are profound (mirroring the national diversity of views about gay marriage, and gay rights in general, as well as about abortion rights). A uniform rule is neither necessary nor appropriate. Yet that is what the *Heller* decision will produce if its rule is

held applicable to the states as well as to the District of Columbia and other federal enclaves.

Heller gives short shrift to the values of federalism, and to the related values of cultural diversity, local preference, and social experimentation. A majority of Americans support gun rights. But if the District of Columbia (or Chicago or New York) wants to ban guns, why should the views of a national majority control? Is that democracy, or is it Rousseau's forced conformity to the "general will"? True, a member of a national majority can be a member of a minority within a local area: gun buffs in Washington, D.C., for example. But a person who is a member of a local minority but a national majority can relocate to a part of the country in which the national majority rules. A resident of Washington can move to northern Virginia. This is not to say that there should be no national rights—that Mississippi should be permitted to stone adulterers, or Rhode Island to ban *The Da Vinci Code.* But the question of whether to nationalize an issue in the name of the Constitution calls for an exercise of judgment; and when the nation is deeply divided over an issue to which the Constitution does not speak with any clarity, and a uniform national policy would override differences in local conditions, nationalization may be premature.

There is a further difference between constitutional interpretations that permit government action and ones that forbid it: only the latter create new business for the federal courts. Conservatives rightly decry the enormous expansion in the federal caseload caused by the aggressive constitutional rulings of liberal justices in the 1960s. But if the new rule declared in *Heller* is applied to the states, we may see a similar result, this time engineered by conservatives; and we will have further confirmation that the Warren Court liberated conservative as well as liberal judges from the constraint of judicial modesty. Every time a gun permit is denied, the disappointed applicant will have a potential constitutional claim litigable in the federal courts.

Justice Scalia was emphatic that the right to possess a gun is not absolute. He sparred with Justice Stephen Breyer (who wrote a separate dissenting opinion) over the standard to be applied to restrictions on gun ownership. All that is clear is that an absolute ban on possessing a pistol is unconstitutional. The other restrictions that a government might want to impose are up for grabs. It may take many years for the dust to settle—many years of lawsuits that our litigious society does not need.

Conservatives rightly believe, moreover, that the efficacy of legally enforceable rights as an engine for social reform is overrated. The effects even of such well known and generally applauded decisions as those invalidating racial segregation of public schools and the malapportionment of state legislatures are uncertain, and may not have been, on balance, beneficial. The only certain effect of the *Heller* decision—for the scholarly literature has yet to reach consensus on the effects of gun-control laws—will be to increase litigation over gun ownership.

IV

I cannot discern any principles in the pattern of the Supreme Court's constitutional interpretations. The absence of principles supports the hypothesis that ideology drives decision in cases in which liberal and conservative values collide. If loose construction produces a conservative limitation on government, most conservatives will support it and most liberals will oppose it; and if it produces a liberal limitation on government, most liberals and conservatives will switch sides. The qualification in "most" is important, though. Scalia has voted to invalidate, on free-speech grounds, laws forbidding the burning of the American flag. That is loose construction—decidedly non-originalist in the narrow sense of his opinion in *Heller*—because burning is not speech; and it is a loose construction that produces a liberal outcome. Breyer concurred in a decision that allowed the Ten Commandments to be exhibited on the grounds of the Texas Capitol; and that was a conservative vote (and the swing vote in the case) by a liberal justice.

Whatever generated these justices' uncharacteristic votes in those two cases, it was not a decision-making method that prevents the exercise of discretion. Both justices employ judicial methodologies that leave them with plenty of running room. In his dissent in the *Zelman* case, Breyer argued that the school voucher system was unconstitutional because there was a "risk" that it could create "a form of religiously based conflict potentially harmful to the Nation's social fabric." In his

dissent in *Heller* he reversed the burden, arguing that the risk that allowing limited gun ownership in the District of Columbia would lead to more death and injury from guns was enough to uphold the District's gun law against constitutional challenge.

Since Stevens devoted most of his dissenting opinion in *Heller* to his own interpretation of the original meaning of the Second Amendment, observers may conclude that the entire Court has now embraced originalism as the canonical method of interpreting the Constitution. But this is not a plausible inference in light of the child-rapist case of just a few weeks earlier (non-murdering rapists of adult women were being executed in the United States as recently as the 1960s), from which Scalia dissented. One supposes that Stevens could not resist meeting the majority on its own ground, since the text and the history (both pre-and post-enactment) of the Second Amendment favor the dissent. Among other things, professional historians were on Stevens's side.

Still, his opinion seems to me too dogmatic (the historical evidence is not as one-sided as his opinion suggests); and it leaves the impression that all that divided the two wings of the Court was a disagreement over the historical record. That was playing into Scalia's hands. The majority (and the dissent as well) was engaged in what is derisively referred to—the derision is richly deserved—as "law office history." Lawyers are advocates for their clients, and judges are advocates for whichever side of the case they have decided to vote for. The judge sends his law clerks scurrying to the library and to the Web for bits and pieces of historical documentation. When the clerks are the numerous and able clerks of Supreme Court justices, enjoying the assistance of the capable staffs of the Supreme Court library and the Library of Congress, and when dozens and sometimes hundreds of amicus curiae briefs have been filed, many bulked out with the fruits of their authors' own law-office historiography, it is a simple matter, especially for a skillful rhetorician such as Scalia, to write a plausible historical defense of his position.

But it was not so simple in *Heller*, and Scalia and his staff labored mightily to produce a long opinion (the majority opinion is almost 25,000 words long) that would convince, or perhaps just overwhelm, the doubters. The range of historical references in the majority opinion is breathtaking, but it is not evidence of disinterested historical inquiry. It is evidence of the ability of well-staffed courts to produce snow jobs. . . .

For more than two centuries, the "right" to private possession of guns, supposedly created by the Second Amendment, had lain dormant. Constitutional rights often lie dormant, spectral subjects of theoretical speculation, until some change in the social environment creates a demand for their vivification and enforcement. But nothing has changed in the social environment to justify giving the Second Amendment a new life discontinuous with its old one: a new wine in a decidedly old wineskin. There is no greater urgency about allowing people to possess guns for self-defense or defense of property today than there was thirty years ago, when the prevalence of violent crime was greater, or for that matter one hundred years ago. Only the membership of the Supreme Court has changed.

If constitutional decisions are to be determined by the balance between liberals and conservatives on the Supreme Court, the fig-leafing that we find in *Heller*—the historicizing glaze on personal values and policy preferences—will continue to be irresistibly tempting to the justices, with their large and tireless staffs and their commitment to a mystique of "objective" interpretation. There is no way to purge political principles from constitutional decision-making, but they do not have to be liberal or conservative principles. A preference for judicial modesty—for less interference by the Supreme Court with the other branches of government—cannot be derived by some logical process from constitutional text or history. It would have to be imposed. It would be a discretionary choice by the justices. But judging from *Heller*, it would be a wise choice. It would go some distance toward de-politicizing the Supreme Court. It would lower the temperature of judicial confirmation hearings, widen the field of selection of justices, and enable the Supreme Court to attend to the many important non-constitutional issues that it is inclined to neglect.

REVIEW AND DISCUSSION QUESTIONS

1. Assess Posner's argument that originalism should have led to a different decision in *Heller.* Explain why you agree or disagree with him that the Second Amendment "whether viewed alone or in light of the concerns that actuated its adoption" creates no right to own a gun for hunting or for personal defense and that it is doubtful that the Amendment even requires that members of the militia be allowed to keep weapons at home.
2. Why is "loose construction" true originalism, according to Posner?
3. Posner favors loose construction when it relaxes rather than tightens the Constitution's grip on the legislative and executive branches. Why? In this connection, compare *Heller* to *Kennedy v. Louisiana*, on the one hand, and to *Kelo* and *Zelman*, on the other hand.
4. Would Posner's position have led him to oppose the decision in *Rochin v. California*? Does Posner do justice to the argument that sometimes tightening Constitutional restrictions on what the government may do is necessary to uphold fundamental rights or basic values?
5. Assess Posner's argument that ideology drives too many Supreme Court decisions. Should the Supreme Court be de-politicized? How important a goal is "judicial modesty"?
6. What does Posner mean by "law office history"?
7. Assess Posner's arguments from the perspective of Ronald Dworkin.

PART THREE

The Nature of Law

CHAPTER 7

Natural Law and Legal Positivism: Classical Perspectives

We begin our reflections on the nature of law by examining two rival, and enduringly important, theories which will be in the background of much that follows: natural law and legal positivism.

The natural law approach is introduced by a general discussion of that approach, followed by selections from Thomas Aquinas's *Summa Theologica* and from William Blackstone's *Commentaries on the Laws of England.* One of the most influential Christian theologians of all time, Aquinas combined in his philosophy elements of classical Greek thought, particularly that of Aristotle, with Christian natural theology. In this selection, he distinguishes different types of law and explains the constraints that exist in human law. William Blackstone, by contrast, was a legal scholar, whose *Commentaries on the Laws of England* outlined the basic structure and guiding principles of English law at the time. Written when there were neither law schools nor legal textbooks, Blackstone's volumes quickly became, and remained for years and years, the single most authoritative guide to English law. As a result of studying Blackstone, most English and American lawyers of the eighteenth and nineteenth centuries implicitly absorbed his natural law philosophy.

In the third reading of this section, John Austin presents the traditional positivist position, which he sees as diametrically opposed to the natural law theory of Aquinas and Blackstone. Although Austin does not deny the existence of divine or natural law, he insists that they are distinct from the law of a state.

Traditional Natural Law Theory

Brian Bix

In this essay, Brian Bix discusses the central tenets of traditional natural law theory. His focus is on the major figures who have defended natural law theory, from Cicero and Aquinas to Grotius. He concludes with a few reflections on the importance of natural law theory to the development of law and especially on its role in American history. Brian Bix is professor of law at the University of Minnesota Law School.

We take it for granted that the laws and legal system under which we live can be criticized on moral grounds: that there are standards against which legal norms can be compared and sometimes found wanting. The standards against which law is judged have sometimes been described as "a (the) higher law." For some, this is meant literally: that there are law-like standards that have been stated in or can be derived from divine revelation, religious texts, a careful study of human nature, or consideration of nature. For others, the reference to "higher law" is meant metaphorically, in which case it at least reflects our mixed intuitions about the moral status of law: on the one hand, that not everything properly enacted as law is binding morally; on the other hand, that the law, as law, does have moral weight. (If it did not, we would not need to point to a "higher law" as a justification for ignoring the requirements of our society's laws.)

"Traditional" natural law theory offers arguments for the existence of a "higher law," elaborations of its content, and analyses of what consequences follow from the existence of a "higher law" (in particular, what response citizens should have to situations where the positive law—the law enacted within particular societies—conflicts with the "higher law").

CICERO

While one can locate a number of passages in ancient Greek writers that express what appear to be natural law positions, including passages in Plato (*Laws, Statesman, Republic*) and Aristotle (*Politics, Nicomachean Ethics*), as well as Sophocles' *Antigone,* the best known ancient formulation of a natural law position was offered by the Roman orator Cicero.

Cicero (*Laws, Republic*) wrote in the first century B.C., and was strongly influenced (as were many Roman writers on law) by the works of the Greek Stoic philosophers (some would go so far as to say that Cicero merely offered an elegant restatement of already established Stoic views). Cicero offered the following characterization of "natural law":

> True law is right reason in agreement with nature; it is of universal application, unchanging and everlasting; it summons to duty by its commands, and averts from wrongdoing by its prohibitions. And it does not lay its commands or prohibitions upon good men in vain, though neither have any effect on the wicked. It is a sin to try to alter this law, nor is it allowable to attempt to repeal any part of it, and it is impossible to abolish it entirely. We cannot be freed from its obligations by senate or people, and we need not look outside ourselves for an expounder or interpreter of it. And there will not be different laws at Rome and at Athens, or different laws now and in the future, but one eternal and unchangeable law will be valid for all nations and all times, and there will be one master and ruler, that is, God, over us all, for he is the author of this law, its promulgator, and its enforcing judge. Whoever is disobedient is fleeing from himself and denying his human nature, and by reason of this very fact he will suffer the worst penalties, even if he escapes what is commonly considered punishment.

In Cicero's discussions of law, we come across most of the themes traditionally associated with traditional natural law theory (though, as might be expected in the first major treatment of a subject, some of the analysis is not always as systematic or as precise as one might want): natural law is unchanging over time and does not differ in different societies; every person has access to the standards of this higher law by use of reason; and only just laws "really deserve [the] name" law, and "in the very definition of the term 'law' there inheres the idea and principle of choosing what is just and true."

Within Cicero's work, and the related remarks of earlier Greek and Roman writers, there was often a certain ambiguity regarding the reference of "natural" in "natural law": it was not always clear whether the standards were "natural" because they derived from "human nature" (our "essence" or "purpose"); because they were accessible by our natural faculties (that is, by human reason or conscience); because they derived from or were expressed in nature, that is, in the physical world about us; or some combination of all three.

As one moves from the classical writers on natural law to the early Church writers, aspects of the theory necessarily change and therefore raise different issues within this approach to morality and law. For example, with classical writers, the source of the higher standards is said to be (or implied as being) inherent in the nature of things. With the

early Church writers, there is a divine being who actively intervenes in human affairs and lays down express commands for all mankind—though this contrast overstates matters somewhat, as the classical writers referred to a (relatively passive) God, and the early Church writers would sometimes refer to rules of nature which express divine will. To the extent that the natural law theorists of the early Church continued to speak of higher standards inherent in human nature or in the nature of things, they also had to face the question of the connection between these standards and divine commands: for example, whether God can change natural law or order something which is contrary to it, a question considered by Ambrose and Augustine (among others) in the time of the early Church and by Francisco Suarez more than a thousand years later.

AQUINAS

The most influential writer within the traditional approach to natural law is undoubtedly Thomas Aquinas (*Summa Theologica*), who wrote in the thirteenth century. The context of Aquinas's approach to law, its occurrence within a larger theological project that offered a systematic moral system, should be kept in mind when comparing his work with more recent theorists.

Aquinas identified four different kinds of law: the eternal law, the natural law, the divine law, and human (positive) law. For present purposes, the important categories are natural law and positive law.

According to Aquinas, (genuine or just) positive law is derived from natural law. This derivation has different aspects. Sometimes natural law dictates what the positive law should be: for example, natural law both requires that there be a prohibition of murder and settles what its content will be. At other times, natural law leaves room for human choice (based on local customs or policy choices). Thus, while natural law would probably require regulation of automobile traffic for the safety of others, the choice of whether driving should be on the left or the right side of the road, and whether the speed limit should be set at 55 miles per hour or 65, are matters for which either choice would probably be compatible with the requirements of natural law. The first form of derivation is like logical deduction; the second Aquinas refers to as the "determination" of general principles ("determination" not in the sense of "finding out," but rather in the sense of making specific or concrete). The theme of different ways in which human (positive) law derives from natural law is carried on by later writers, including Sir William Blackstone (*Commentaries on the Laws of England,* Vol. I [1765]), and, in modern times, John Finnis.

As for citizens, the question is what their obligations are regarding just and unjust laws. According to Aquinas, positive laws which are just "have the power of binding in conscience." A just law is one which is consistent with the requirements of natural law—that is, it is "ordered to the common good," the lawgiver has not exceeded its authority, and the law's burdens are imposed on citizens fairly. Failure with respect to any of those three criteria, Aquinas asserts, makes a law unjust; but what is the citizen's obligation in regard to an unjust law? The short answer is that there is no obligation to obey that law. However, a longer answer is warranted, given the amount of attention this question usually gets in discussions of natural law theory in general and of Aquinas in particular.

The phrase *lex iniusta non est lex* ("an unjust law is not law") is often ascribed to Aquinas, and is sometimes given as a summation of his position and the (traditional) natural law position in general. While Aquinas never used the exact phrase above, one can find similar expressions: "every human law has just so much of the nature of law, as it is derived from the law of nature. But if in any point it deflects from the law of nature, it is no longer a law but a perversion of law"; and "[unjust laws] are acts of violence rather than laws; because, as Augustine says, a law that is not just, seems to be no law at all." (One also finds similar statements by Plato, Aristotle, Cicero, and Augustine—though, with the exception of Cicero's, these statements are not part of a systematic discussion of the nature of law.)

Questions have been raised regarding the significance of the phrase. What does it mean to say that an apparently valid law is "not law," "a perversion of law," or "an act of violence rather than a law"? Statements of this form have been offered and

interpreted in one of two ways. First, one can mean that an immoral law is not valid law at all. The nineteenth-century English jurist John Austin (*The Province of Jurisprudence Determined* [1832]) interpreted statements by Sir William Blackstone (for example, "no human laws are of any validity, if contrary to [the law of nature]") in this manner, and pointed out that such analyses of validity are of little value. Austin wrote,

> Suppose an act innocuous, or positively beneficial, be prohibited by the sovereign under the penalty of death; if I commit this act, I shall be tried and condemned, and if I object to the sentence, that it is contrary to the law of God . . . the Court of Justice will demonstrate the inconclusiveness of my reasoning by hanging me up, in pursuance of the law of which I have impugned the validity.

Though one must add that we should not conflate questions of power with questions of validity—for a corrupt legal system might punish someone even if shown that the putative law was invalid under the system's own procedural requirements—we understand the distinction between validity under the system's rules and the moral worth of the enactment in question.

A more reasonable interpretation of statements like "an unjust law is no law at all" is that unjust laws are not laws "in the fullest sense." As we might say of some professional, who had the necessary degrees and credentials, but seemed nonetheless to lack the necessary ability or judgment: "she's no lawyer" or "he's no doctor." This only indicates that we do not think that the title in this case carries with it all the implications it usually does. Similarly, to say that an unjust law is "not really law" may only be to point out that it does not carry the same moral force or offer the same reasons for action as laws consistent with "higher law." This is almost certainly the sense in which Aquinas made his remarks, and the probable interpretation for nearly all proponents of the position. However, this interpretation leaves the statement as clearly right as under the prior (Austinian) interpretation it was clearly wrong. One wonders why such declarations have, historically, been so controversial.

To say that an unjust law is not law in the fullest sense is usually intended not as a simple declaration, but as the first step of a further argument. For example: "this law is unjust; it is not law in the fullest sense, and therefore citizens can in good conscience act as if it was never enacted; that is, they should feel free to disobey it." This is a common understanding of the idea that an unjust law is no law at all, but it expresses a conclusion that is controversial.

There are often moral reasons for obeying even an unjust law: for example, if the law is part of a generally just legal system, and public disobedience of the law might undermine the system, there is a moral reason for at least minimal, public obedience to the unjust law. This is Aquinas's position (he stated that a citizen is not bound to obey "a law which imposes an unjust burden on its subjects" if the law "can be resisted without scandal or greater harm"), and it has been articulated at greater length by later natural law theorists (for example, by John Finnis).

Finally, it should be noted that the proper interpretation of certain basic aspects of Aquinas's work remains in dispute. For example, there is debate within the modern literature regarding whether Aquinas believed moral norms could be derived directly from knowledge of human nature or experience of natural inclinations, or whether they are the product of practical understanding and reasoning by way of reflection on one's experience and observations.

NATURAL LAW IN EARLY MODERN EUROPE

In the period of the Renaissance and beyond, discussions about natural law were tied in with other issues: assertions about natural law were often the basis of or part of the argument for individual rights and limitations on government. Such discussions were also often the groundwork offered for principles of international law. Hugo Grotius and Samuel Pufendorf (writing in the early and late seventeenth century, respectively) were prominent examples of theorists whose writings on natural law had significance in both debates. (Grotius and Pufendorf, along with other prominent seventeenth-century theorists, Francisco Suarez, Thomas Hobbes, and John Locke, were also central in developing the concept of individual rights in the modern sense of that term.)

A further significance of Grotius' work was its express assertion that natural law, the higher law against which the actions of nations, law-makers, and citizens could be judged, did not require the existence of God for its validity. (However, one can find hints of such a separation of natural law from a divine being at least as far back as the fourteenth-century writings of Gregory of Rimini.) From that time to the present, an increasingly large portion of the writing on questions of natural law (and the related idea of "natural rights") was secular in tone and purpose, usually referring to "the requirements of reason" rather than divine command, purpose, will, or wisdom.

PERSPECTIVE

It is normally a mistake to try to evaluate the discussions of writers from distant times with the perspective of modern analytical jurisprudence.

Cicero, Aquinas, and Grotius were not concerned with a social–scientific-style analysis of law, as the modern advocates of legal positivism could be said to be. These theorists were concerned with what legislators, citizens, and governments ought to do, or could do in good conscience. It is not that these writers (and their followers) never asked questions like "what is law?" However, they were asking the questions as a starting point for an ethical inquiry, and therefore one should not be too quick in comparing their answers with those in similar-sounding discussions by recent writers, who see themselves as participating in a conceptual or sociological task.

Natural law has, from time to time and with varying degrees of importance, escaped the confines of theory to influence directly the standards created and applied by officials. For example, natural law (or standards and reasoning that appear similar to natural law, but which are characterized as "substantive due process," "natural justice," or simply "reason") has been offered as the source of legal standards for international law, centuries of development in the English common law, and certain aspects of United States constitutional law. Natural law also appears to have played a significant role in American history, where its reasoning, or at least its rhetoric, has been prominent (among other places) in the Declaration of Independence, the Abolition (anti-slavery) movement, and parts of the modern Civil Rights movement.

REVIEW AND DISCUSSION QUESTIONS

1. What are the central tenets of traditional natural law theory, as exemplified in the writings of Cicero?
2. Describe the four types of law, according to Aquinas.
3. What does Bix think is the most reasonable interpretation of the idea that an unjust law is not a law?
4. What was Grotius' contribution to traditional natural law theory?
5. Describe how Bix thinks natural law has affected American history.
6. "Natural law theory merely confuses what the law is with what it ought to be." Discuss that claim in light of Bix's article.

Summa Theologica

Thomas Aquinas

Saint Thomas Aquinas was born in Italy in 1224. He was of noble birth and was able to enter a Dominican order and spend his life as a teacher and scholar in Paris, Cologne, Rome, and Naples. With the rediscovery of ancient Greek philosophy (in particular, Aristotle's writings), the Roman Catholic Church was faced with the problem of reconciling its theology with the teachings of the Greeks. Thomas Aquinas undertook this task, developing an all-embracing philosophical system that sought a synthesis of faith and reason. He died at the age of fifty, and in 1879, the Roman Catholic Church declared his writings to be its official teaching.

In this brief selection from his classic work, *Summa Theologica,* Aquinas first provides a definition of law and then describes the four types of law along with the interrelationships among them. As background, Aquinas assumes with Aristotle that each creature has its own natural purpose or end and that fulfilling that purpose defines its good. Just as one cannot understand a knife or an eye without appreciating its purpose, so too for Aristotle it is impossible to understand a legal or a political system without first thinking about the natural good of human beings, as well as the purpose of law itself.

QUESTION 90: ON THE ESSENCE OF LAW

Whether Law Is Always Directed to th Common Good?

As we have stated above, law belongs to that which is a principle of human acts, because it is their rule and measure. Now as reason is a principle of human acts, so in reason itself there is something which is the principle in respect of all the rest. Hence to this principle chiefly and mainly law must needs be referred. Now the first principle in practical matters, which are the object of the practical reason, is the last end: and the last end of human life is happiness or beatitude, as we have stated above. Consequently, law must needs concern itself mainly with the order that is in beatitude. Moreover, since every part is ordained to the whole as the imperfect to the perfect, and since one man is a part of the perfect community, law must needs concern itself properly with the order directed to universal happiness. Therefore the Philosopher [Aristotle], in the above definition of legal matters, mentions both happiness and the body politic, since he says that we call those legal matters *just which are adapted to produce and preserve happiness and its parts for the body politic.* For the state is a perfect community, as he says in *Politics* i.

Now, in every genus, that which belongs to it chiefly is the principle of the others, and the others belong to that genus according to some order towards that thing. Thus fire, which is chief among hot things, is the cause of heat in mixed bodies, and these are said to be hot in so far as they have a share of fire. Consequently, since law is chiefly ordained to the common good, any other precept in regard to some individual work must needs be devoid of the nature of a law, save in so far as it regards the common good. Therefore every law is ordained to the common good. . . .

Whether the Reason of Any Man Is Competent to Make Laws?

. . . A private person cannot lead another to virtue efficaciously; for he can only advise, and if his advice be not taken, it has no coercive power, such as the law should have, in order to prove an efficacious inducement to virtue, as the Philosopher says. But this coercive power is vested in the whole people or in some public personage, to whom it belongs to inflict penalties, as we shall state further on. Therefore the framing of laws belongs to him alone.

. . . As one man is a part of the household, so a household is a part of the state; and the state is a perfect community, according to *Politics* i. Therefore, just as the good of one man is not the last end, but is ordained to the common good, so too the good of one household is ordained to the good of a single state, which is a perfect community. Consequently,

he that governs a family can indeed make certain commands or ordinances, but not such as to have properly the nature of law.

Whether Promulgation Is Essential to Law?

As was stated above, a law is imposed on others as a rule and measure. Now a rule or measure is imposed by being applied to those who are to be ruled and measured by it. Therefore, in order that a law obtain the binding force which is proper to a law, it must needs [be] applied to the men who have to be ruled by it. But such application is made by its being made known to them by promulgation. Therefore promulgation is necessary for law to obtain its force.

Thus, . . . the definition of law may be gathered. Law is nothing else than an ordinance of reason for the common good, promulgated by him who has the care of the community.

. . . The natural law is promulgated by the very fact that God instilled it into man's mind so as to be known by him naturally.

. . . Those who are not present when a law [is] promulgated are bound to observe the law, in so far as it is made known or can be made known to them by others, after it has been promulgated. . . .

QUESTION 91: ON THE VARIOUS KINDS OF LAW

Whether There Is an Eternal Law?

Objection: It would seem that there is no eternal law. For every law is imposed on someone. But there was not someone from eternity on whom a law could be imposed, since God alone was from eternity. Therefore no law is eternal. . . .

On the contrary, Augustine says: *That Law which is the Supreme Reason cannot be understood to be otherwise than unchangeable and eternal.*

I answer that . . . law is nothing else but a dictate of practical reason emanating from the ruler who governs a perfect community. Now it is evident, granted that the world is ruled by divine providence, . . . that the whole community of the universe is governed by the divine reason. Therefore the very notion of the government of things in God, the ruler of the universe, has the nature of a law. And since the divine reason's conception of things is not subject to time, but is eternal, according to *Prov.* viii. 23, therefore it is that this kind of law must be called eternal. . . .

Whether There Is in Us a Natural Law?

. . . Law, being a rule and measure, can be in a person in two ways: in one way, as in him that rules and measures; in another way, as in that which is ruled and measured, since a thing is ruled and measured in so far as it partakes of the rule or measure. Therefore, since all things subject to divine providence are ruled and measured by the eternal law, as was stated above, it is evident that all things partake in some way in the eternal law, in so far as, namely, from its being imprinted on them, they derive their respective inclinations to their proper acts and ends. Now among all others, the rational creature is subject to divine providence in a more excellent way, in so far as it itself partakes of a share of providence, by being provident both for itself and for others. Therefore it has a share of the eternal reason, whereby it has a natural inclination to its proper act and end; and this participation of the eternal law in the rational creature is called the natural law. . . . [T]he light of natural reason, whereby we discern what is good and what is evil, which is the function of the natural law, is nothing else than an imprint on us of the divine light. It is therefore evident that the natural law is nothing else than the rational creature's participation of the eternal law. . . .

Whether There Is a Human Law?

. . . A law is a dictate of the practical reason. Now it is to be observed that the same procedure takes place in the practical and in the speculative reason, for each proceeds from principles to conclusions, as was stated above. Accordingly, we conclude that, just as in the speculative reason, from naturally known indemonstrable principles we draw the conclusions of the various sciences, the knowledge of which is not imparted to us by nature, but acquired by the efforts of reason, so too it is that from the precepts of the natural law, as from common and indemonstrable principles, the human reason needs to proceed to the more particular determination of

certain matters. These particular determinations, devised by human reason, are called human laws, provided that the other essential conditions of law be observed, as was stated above.

. . . The human reason cannot have a full participation of the dictate of the divine reason, but according to its own mode, and imperfectly. Consequently, just as on the part of the speculative reason, by a natural participation of divine wisdom, there is in us the knowledge of certain common principles, but not a proper knowledge of each single truth, such as that contained in the divine wisdom, so, too, on the part of the practical reason, man has a natural participation of the eternal law, according to certain common principles, but not as regards the particular determinations of individual cases, which are, however, contained in the eternal law. Hence the need for human reason to proceed further to sanction them by law. . . .

Whether There Was Any Need for a Divine Law?

. . . Besides the natural and the human law it was necessary for the directing of human conduct to have a divine law. And this for four reasons. First, because it is by law that man is directed how to perform his proper acts in view of his last end. Now if man were ordained to no other end than that which is proportionate to his natural ability, there would be no need for man to have any further direction, on the part of his reason, in addition to the natural law and humanly devised law which is derived from it. But since man is ordained to an end of eternal happiness which exceeds man's natural ability, as we have stated above, therefore it was necessary that, in addition to the natural and the human law, man should be directed to his end by a law given by God.

Secondly, because, by reason, of the uncertainty of human judgment, especially on contingent and particular matters, different people form different judgments on human acts; whence also different and contrary laws result. In order, therefore, that man may know without any doubt what he ought to do and what he ought to avoid, it was necessary for man to be directed in his proper acts by a law given by God, for it is certain that such a law cannot err.

Thirdly, because man can make laws in those matters of which he is competent to judge. But man is not competent to judge of interior movements, that are hidden, but only of exterior acts which are observable; and yet for the perfection of virtue it is necessary for man to conduct himself rightly in both kinds of acts. Consequently, human law could not sufficiently curb and direct interior acts, and it was necessary for this purpose that a divine law should supervene.

Fourthly, because, as Augustine says, human law cannot punish or forbid all evil deeds, since, while aiming at doing away with all evils, it would do away with many good things, and would hinder the advance of the common good, which is necessary for human living. In order, therefore, that no evil might remain unforbidden and unpunished, it was necessary for the divine law to supervene, whereby all sins are forbidden. . . .

QUESTION 92: ON THE EFFECTS OF LAW

Whether It Is an Effect of Law to Make Men Good?

. . . A law is nothing else than a dictate of reason in the ruler by whom his subjects are governed. Now the virtue of any being that is a subject consists in its being well subordinated to that by which it is regulated; and thus we see that the virtue of the irascible and concupiscible powers consists in their being obedient to reason. In the same way, *the virtue of every subject consists in his being well subjected to his ruler,* as the Philosopher says. But every law aims at being obeyed by those who are subject to it. Consequently it is evident that the proper effect of law is to lead its subjects to their proper virtue; and since virtue is *that which makes its subject good,* it follows that the proper effect of law is to make those, to whom it is given, good, either absolutely or in some particular respect. For if the intention of the lawgiver is fixed on a true good, which is the common good regulated according to divine justice, it follows that the effect of law is to make men good absolutely. If, however, the intention of the lawgiver is fixed on that which is not good absolutely, but useful or pleasurable to himself, or in opposition to divine justice, then law does not make men good absolutely, but in a relative way, namely, in relation to that particular government. In this way good is found even in things that are bad of themselves. Thus a man is called a good robber, because he works in a way that is adapted to his end.

. . . A tyrannical law, through not being according to reason, is not a law, absolutely speaking, but rather a perversion of law; and yet in so far as it is something in the nature of a law, its aim is that the citizens be good. For it has the nature of law only in so far as it is an ordinance made by a superior to his subjects, and aims at being obeyed by them; and this is to make them good, not absolutely, but with respect to that particular government. . . .

QUESTION 93: THE ETERNAL LAW

. . . Just as in every artificer there pre-exists an exemplar of the things that are made by his art, so too in every governor there must pre-exist the exemplar of the order of those things that are to be done by those who are subject to his government. And just as the exemplar of the things yet to be made by an art is called the art or model of the products of that art, so, too, the exemplar in him who governs the acts of his subjects bears the character of a law, provided the other conditions be present which we have mentioned above as belonging to the nature of law. Now God, by His wisdom, is the Creator of all things, in relation to which He stands as the artificer to the products of His art, as was stated in the First Part. Moreover, He governs all the acts and movements that are to be found in each single creature, as was also stated in the First Part. Therefore, just as the exemplar of the divine wisdom, inasmuch as all things are created by it, has the character of an art, a model or an idea, so the exemplar of divine wisdom, as moving all things to their due end, bears the character of law. Accordingly, the eternal law is nothing else than the exemplar of divine wisdom. As directing all actions and movements. . . .

Since, then, the eternal law is the plan of government in the Chief Governor, all the plans of government in the inferior governors must be derived from the eternal law. But these plans of inferior governors are all the other laws which are in addition to the eternal law. Therefore all laws, in so far as they partake of right reason, are derived from the eternal law. Hence Augustine says that *in temporal law there is nothing just and lawful but what man has drawn from the eternal law.*

. . . Human law has the nature of law in so far as it partakes of right reason; and it is clear that, in this respect, it is derived from the eternal law. But in so far as it deviates from reason, it is called an unjust law, and has the nature, not of law, but of violence. Nevertheless, even an unjust law, in so far as it retains some appearance of law, through being framed by one who is in power, is derived from the eternal law; for all power is from the Lord God, according to *Rom* xiii, I.

. . . Human law is said to permit certain things, not as approving of them, but as being unable to direct them. And many things are directed by the divine law, which human law is unable to direct, because more things are subject to a higher than a lower cause. Hence the very fact that human law does not concern itself with matters it cannot direct comes under the ordination of the eternal law. It would be different, were human law to sanction what the eternal law condemns. Consequently, it does not follow that human law is not derived from the eternal law; what follows is rather that it is not on a perfect equality with it. . . .

QUESTION 94: THE NATURAL LAW

Whether the Natural Law Contains Several Precepts or Only One?

. . . The precepts of the natural law are to the practical reason what the first principles of demonstrations are to the speculative reason, because both are self-evident principles. Now a thing is said to be self-evident in two ways: first, in itself; secondly, in relation to us. Any proposition is said to be self-evident in itself, if its predicate is contained in the notion of the subject; even though it may happen that to one who does not know the definition of the subject, such a proposition is not self-evident. For instance, this proposition, *Man is a rational being,* is in its very nature, self-evident, since he who says *man,* says *a rational being;* and yet to one who does not know what a man is, this proposition is not self-evident. Hence it is that, as Boethius says, certain axioms or propositions are universally self-evident to all; and such are the propositions whose terms are known to all, as, *Every whole is greater than its part,* and, *Things equal to one and the same*

are equal to one another. But some propositions are self-evident only to the wise, who understand the meaning of the terms of such propositions. Thus to one who understands that an angel is not a body, it is self-evident that an angel is not circumscriptively in a place. But this is not evident to the unlearned, for they cannot grasp it.

. . . Now as *being* is the first thing that falls under the apprehension absolutely, so *good* is the first thing that falls under the apprehension of the practical reason, which is directed to action (since every agent acts for an end, which has the nature of good). Consequently, the first principle in the practical reason is one founded on the nature of good, viz., that *good is that which all things seek after.* Hence this is the first precept of law, that *good is to be done and promoted, and evil is to be avoided.* All other precepts of the natural law are based upon this; so that all the things which the practical reason naturally apprehends as man's good belong to the precepts of the natural law under the form of things to be done or avoided.

Since, however, good has the nature of an end, and evil, the nature of the contrary, hence it is that all those things to which man has a natural inclination are naturally apprehended by reason as being good, and consequently as objects of pursuit, and their contraries as evil, and objects of avoidance. Therefore, the order of the precepts of the natural law is according to the order of natural inclinations. For there is in man, first of all, an inclination to good in accordance with the nature which he has in common with all substances, inasmuch, namely, as every substance seeks the preservation of its own being, according to its nature; and by reason of this inclination, whatever is a means of preserving human life, and of warding off its obstacles, belongs to the natural law. Secondly, there is in man an inclination to things that pertain to him more specially, according to that nature which he has in common with other animals; and in virtue of this inclination, those things are said to belong to the natural law *which nature has taught to all animals,* such as sexual intercourse, the education of offspring and so forth. Thirdly, there is in man an inclination to good according to the nature of his reason, which nature is proper to him. Thus man has a natural inclination to know the truth about God, and to live in society; and in this respect, whatever pertains to this inclination belongs to the natural law: *e.g.,* to shun ignorance, to avoid offending those among whom one has to live, and other such things regarding the above inclination.

. . . All these precepts of the law of nature have the character of one natural law, inasmuch as they flow from one first precept.

. . . All the inclinations of any parts whatsoever of human nature, *e.g.,* of the concupiscible and irascible parts, in so far as they are ruled by reason, belong to the natural law, and are reduced to one first precept, as was stated above. And thus the precepts of the natural law are many in themselves, but they are based on one common foundation. . . .

QUESTION 95: HUMAN LAW

Whether It Was Useful for Laws to Be Framed by Men?

. . . As we have stated above, man has a natural aptitude for virtue; but the perfection of virtue must be acquired by man by means of some kind of training. . . . Now it is difficult to see how man could suffice for himself in the matter of this training, since the perfection of virtue consists chiefly in withdrawing man from undue pleasures, to which above all man is inclined, and especially the young, who are more capable of being trained. Consequently a man needs to receive this training from another, whereby to arrive at the perfection of virtue. And as to those young people who are inclined to acts of virtue by their good natural disposition, or by custom, or rather by the give of God, paternal training suffices, which is by admonitions. But since some are found to be dissolute and prone to vice, and not easily amenable to words, it was necessary for such to be restrained from evil by force and fear, in order that, at least, they might desist from evil-doing, and leave others in peace, and that they themselves, by being habituated in this way, might be brought to do willingly what hitherto they did from fear, and thus become virtuous. Now this kind of training, which compels through fear of punishment, is the discipline of laws. Therefore, in order that man might have peace and virtue, it was necessary for laws to be framed. . . .

Whether Every Human Law Is Derived from the Natural Law?

. . . As Augustine says, *that which is not just seems to be no law at all.* Hence the force of a law depends on the extent of its justice. Now in human affairs a thing is said to be just from being right, according to the rule of reason. But the first rule of reason is the law of nature, as is clear from what has been stated above. Consequently, every human law has just so much of the nature of law as it is derived from the law of nature. But if in any point it departs from the law of nature, it is no longer a law but a perversion of law.

REVIEW AND DISCUSSION QUESTIONS

1. Describe the nature or essence of law, according to Aquinas, indicating its connection with the common good.
2. How does he define "law"?
3. Describe briefly each of the types of law he distinguishes.
4. "Lawmaking is purposive and rational." Explain, indicating what Aquinas thinks its purpose is and why law cannot be arbitrary.

Blackstone's Commentaries

William Blackstone

Born in 1723, William Blackstone studied at Oxford before becoming a barrister. After several years of legal practice, he decided to concentrate on the academic study of law and in 1753 began giving the first set of lectures on English law ever delivered in a university. These lectures became his famous *Commentaries on the Laws of England* (4 volumes, 1765–69), which endeavored to give order and form to English common law. Perhaps the most thorough treatment of English law ever produced by one man, Blackstone's *Commentaries* shaped the legal thinking of generations of lawyers and judges in both Britain and the United States. He maintained that human law derives its validity from conformity to natural and divine law. "No human laws," he wrote, "are of any validity if contrary to this."

OF THE NATURE OF LAWS IN GENERAL

Law, in its most general and comprehensive sense, signifies a rule of action, and is applied indiscriminately to all kinds of action, whether animate or inanimate, rational or irrational. Thus we say, the laws of motion, of gravitation, of optics, or mechanics, as well as the laws of nature and of nations. And it is that rule of action which is prescribed by some superior, and which the inferior is bound to obey.

Thus, when the Supreme Being formed the universe, and created matter out of nothing, he impressed certain principles upon that matter, from which it can never depart, and without which it would cease to be. When he put that matter into motion, he established certain laws of motion, to which all movable bodies must conform. And, to descend from the greatest operations to the smallest, when a workman forms a clock, or other piece of mechanism, he establishes, at his own pleasure, certain arbitrary laws for its direction, as that the hand shall describe a space in a given time to it, which law as long as the work conforms, so long, continues in perfection, and answers the end of its formation. If we farther advance, from mere inactive matter to vegetable and animal life, we shall find them still governed by laws, more numerous

indeed, but equally fixed and invariable. The whole progress of plants, from the seed to the root, and from thence to the seed again; the method of animal nutrition, digestion, secretion, and all other branches of vital economy; are not left to chance, or the will of the creature itself, but are performed in a wondrous involuntary manner, and guided by unerring rules laid down by the great Creator.

This, then is the general signification of law: a rule of action dictated by some superior being, and, in those creatures that have neither the power to think, nor to will, such laws must be invariably obeyed, so long as the creature itself subsists, for its existence depends on that obedience. But laws, in their more confined sense, and in which it is our present business to consider them, denote the rules, not of action in general, but of action or conduct; that is the precepts by which man, the noblest of all sublunary beings, a creature endowed with both reason and freewill, is commanded to make use of those faculties in the general regulation of his behaviour.

Man, considered as a creature, must necessarily be subject to the laws of his Creator, for he is entirely a dependent being. A being, independent of any other, has no rule to pursue but such as he describes to himself; but a state of dependence will inevitably oblige the inferior to take the will of him on whom he depends as the rule of his conduct; not, indeed, in every particular, but in all those points wherein his dependence consists. This principle, therefore, has more or less extent and effect, in proportion as the superiority of the one and the dependence of the other is greater or less, absolute or limited. And consequently, as man depends absolutely upon his Maker for every thing, it is necessary that he should, in all points, conform to his maker's will.

LAW OF NATURE

This will of his Maker is called the law of nature. For as God, when he created matter, and endowed it with a principle of mobility established certain rules for the perpetual direction of that motion, so, when he created man, and endued him with free will to conduct himself in all parts of life, he laid down certain immutable laws of human nature, whereby that free will is in some degree regulated and restrained, and gave him also the faculty of reason to discover the purport of those laws.

Considering the Creator only, as a being of infinite power, he was able unquestionably to have prescribed whatever laws he pleased to his creature, man, however unjust or severe. But, as he is also a being of infinite wisdom, he has laid down only such laws as were founded in those relations of justice that existed in the nature of things antecedent to any positive precept. These are the eternal immutable laws of good and evil, to which the Creator himself, in all his dispensations, conforms; and which he has enabled human reason to discover, so far as they are necessary for the conduct of human actions. Such, among others, are these principles: that we should live honestly, should hurt nobody, and should render to every one his due; to which these three general precepts Justinian has reduced the whole doctrine of law. . . .

HUMAN LAW MUST NOT CONTRAVENE NATURE

This law of nature, being coeval with mankind, and dictated by God himself, is of course superior in obligation to any other. It is binding over all the globe, in all countries, and at all times: no human laws are of any validity, if contrary to this; and such of them as are valid derive all their force, and all their authority, mediately or immediately, from this original.

But in order to apply this to the particular exigencies of each individual, it is still necessary to have recourse to reason, whose office it is to discover, as was before observed, what the law of nature directs in every circumstance of life, by considering what method will tend the most effectually to our own substantial happiness. And if our reason were always, as in our first ancestor before his transgression, clear and perfect, unruffled by passions, unclouded by prejudice, unimpaired by disease or intemperance, the task would be pleasant and easy; we should need no other guide but this. But every man now finds the contrary in his own experience; that his reason is corrupt, and his understanding full of ignorance and effort.

. . . Undoubtedly the revealed law [found in Holy Scripture] is of infinitely more authenticity than that moral system which is framed by ethical writers, and

denominated the natural law; because one is the law of nature, expressly declared so to be by God himself; the other is only what, by the assistance of human reason, we imagine to be that law. If we could be as certain of the latter as we are of the former, both would have an equal authority; but, till then, they can never be put in any competition together.

Upon these two foundations, the law of nature and the law of revelation, depend all human laws; that is to say, no human laws should be suffered to contradict these. There are, it is true, a great number of indifferent points in which both the divine law and the natural leave a man at his own liberty, but which are found necessary, for the benefit of society, to be restrained within certain limits. And herein it is that human laws have their greatest force and efficacy; for, with regard to such points as are not indifferent, human laws are only declaratory of, and act in subordination to, the former. To instance in the case of murder: this is expressly forbidden by the divine, and demonstrably by the natural law; and, from these prohibitions, arises the true unlawfulness of this crime. Those human laws that annex a punishment to it do not at all increase its moral guilt, or super-add any fresh obligation, *in foro conscientiae* (in the court of conscience), to abstain from its perpetration. Nay, if any human law should allow or enjoin us to commit it, we are bound to transgress that human law, or else we must offend both the natural and the divine. But, with regard to matters that are in themselves indifferent, and are not commanded or forbidden by those superior laws—such, for instance, as exporting of wool into foreign countries—here the inferior legislature has scope and opportunity to interpose, and to make that action unlawful which before was not so.

THE STATE OF NATURE

If man were to live in a state of nature, unconnected with other individuals, there would be no occasion for any other laws than the law of nature and the law of God. Neither could any other law possibly exist: for a law always supposes some superior who is to make it; and, in a state of nature, we are all equal, without any other superior but Him who is the author of our being. But man was formed for society; and, as is demonstrated by the writers on this subject, is neither capable of living alone, nor indeed has the courage to do it. . . .

MUNICIPAL LAW

. . . I proceeded to treat more fully of the principal subject of this section, municipal or civil law; that is, the rule by which particular districts, communities or nations, are governed; being thus defined by Justinian, "*jus civile est quod quisque sibi populus constituit*" (the civil law is that which every nation has established for its own government). I call it municipal law, in compliance with common speech; for, though strictly that expression denotes the particular customs of one single municipium or free town, yet it may with sufficient propriety be applied to any one state or nation, which is governed by the same laws and customs.

DEFINITION

Municipal law, thus understood, is properly defined to be "a rule of civil conduct prescribed by the supreme power in a state, commanding what is right and prohibiting what is wrong." Let us endeavor to explain its several properties, as they arise out of this definition. And, first it is a rule: not a transient sudden order from a superior to or concerning a particular person; but something permanent, uniform, and universal. . . . It is also called a rule, to distinguish it from advice or counsel, which we are at liberty to follow or not, as we see proper, and to judge upon the reasonableness or unreasonableness of the thing advised: whereas our obedience to the law depends not upon our approbation, but upon the maker's will. Counsel is only matter of persuasion, law is matter of injunction; counsel acts only upon the willing, law upon the unwilling also.

It is also called a rule to distinguish it from a compact or agreement; for a compact is a promise proceeding from us, law is a command directed to us. The language of a compact is, "I will, or will not, do this"; that of a law is, "thou shalt, or shalt not, do it." It is true there is an obligation which a compact carries with it, equal in point of conscience to that of a law; but then the original of the obligation is

different. In compacts, we ourselves determine and promise what shall be done, before we are obliged to do it; in laws, we are obliged to act without ourselves determining or promising anything at all. Upon these accounts law is defined to be "a rule."

MUNICIPAL LAW (AGAIN)

Municipal law is "a rule of civil conduct." This distinguishes municipal law from the natural, or revealed; the former of which is the rule of moral conduct, and the latter not only the rule of moral conduct, but also the rule of faith. These regard man as a creature, and point out his duty to God, to himself, and to his neighbour, considered in the light of an individual. But municipal or civil law regards him also as a citizen, and bound to other duties towards his neighbour than those of mere nature and religion; duties, which he has engaged in by enjoying the benefits of the common union; and which amount to no more than that he do contribute, on his part, to the subsistence and peace of the society.

It is likewise "a rule prescribed." Because a bare resolution, confined in the breast of the legislator, without manifesting itself by some external sign, can never be properly a law. It is requisite that this resolution be notified to the people who are to obey it. But the manner in which this notification is to be made, is matter of very great indifference. It may be notified by universal tradition and long practice, which supposes a previous publication, and is the case of the common law of England. It may be notified *via voce* (by word of mouth), by officers appointed to be publicly read in churches and other as it may lastly be notified by writing, printing, or the like; which is the general course taken with all our acts of Parliament. Yet, whatever way is made use of, it is incumbent on the promulgators to do it in the most public and perspicuous manner, not like Caligula who . . . wrote his laws in a very small character, and hung them upon high pillars, the more effectually to ensnare the people. There is still a more unreasonable method than this, which is called making of laws *ex post facto* (after the fact); when after an action (indifferent in itself) is committed, the legislature then for the first time declares it to have been a crime, and inflicts a punishment upon the person who has committed it. Here it is impossible that the party could foresee that an action innocent when it was done, should be afterwards converted to guilt by a subsequent law; he had therefore no cause to abstain from it; and all punishment for not abstaining must of consequence be cruel and unjust. All laws should be therefore made to commence *in futuro* (at a future period), and be notified before their commencement; which is implied in the term "prescribed." But when this rule is in the usual manner notified or prescribed, it is then the subject's business to be thoroughly acquainted therewith; for if ignorance, of what he might know, were admitted as a legitimate excuse the laws would be of no effect, but might always be eluded with impunity. . . .

DEFINITION CONTINUED—RIGHT AND WRONG CLAUSE

From what has been advanced, the truth of the former branch of our definition, is (I trust) sufficiently evident; that "municipal law is a rule of civil conduct prescribed by the supreme power in a state." I proceed now to the latter branch of it; that it is a rule so prescribed, "commanding what is right, and prohibiting what is wrong."

Now in order to do this completely, it is first of all necessary that the boundaries of right and wrong be established and ascertained by law. And when this is once done, it will follow of course that it is likewise the business of the law, considered as a rule of civil conduct, to enforce these rights, and to restrain or redress these wrongs. It remains therefore only to consider in which manner the law is said to ascertain the boundaries of right and wrong; and the methods which it takes to command the one and prohibit the other.

SPECIES OF LAWS

For this purpose every law may be said to consist of several parts; one declaratory; whereby the rights to be observed, and the wrongs to be eschewed, are clearly defined and laid down: another, directory; whereby the subject is instructed and enjoined to

observe those rights, and to abstain from the commission of those wrongs: a remedial; whereby a method is pointed out to recover a man's private rights, or redress his private wrongs: to which may be added a fourth, usually termed the sanction or vindicatory branch of the law; whereby it is signified what evil or penalty shall be incurred by such as commit any public wrongs, and transgress or neglect their duty. . . .

ACTS *MALUM PROHIBITUM ET MALUM IN SE* (CRIMES WRONG BECAUSE FORBIDDEN AS DISTINCT FROM WRONG IN THEMSELVES)

It is true . . . that human laws are binding upon men's consciences. But if that were the only or most forcible obligation, the good only would regard the laws, and the bad would set them at defiance. And, true as this principle is, it must still be understood with some restriction. It holds, I apprehend, as to rights; and that, when the law has determined the field to belong to [someone], it is matter of conscience no longer to withhold or to invade it. So also in regard to natural duties, and such offenses as are *mala in se* (wrong in themselves); here we are bound in conscience; because we are bound by superior laws, before those human laws were in being, to perform the one and abstain from the other. But in relation to those laws, which enjoin only positive duties, and forbid only such things as are not *mala in se* (wrong in themselves), but *mala prohibita* (crimes, because forbidden), merely, without any intermixture of moral guilt, annexing a penalty to non-compliance, here I apprehend conscience is no farther concerned, than by directing a submission to the penalty, in case of our breach of those laws: for otherwise the multitude of penal laws in a state would not only be looked upon as an impolitic, but would also be a very wicked thing; if every such law were a snare for the conscience of the subject. But in these cases the alternative is offered to every man; "either abstain from this, or submit to such a penalty": and his conscience will be clear, which ever side of the alternative he thinks proper to embrace. Thus, by the statutes for preserving the game, a penalty is denounced against every unqualified person that kills a hare, and against every person who possesses a partridge in August. And so, too, by other statutes, pecuniary penalties are inflicted for exercising trades without serving an apprenticeship thereto, for not burying the dead in woolen, for not performing the statute-work on the public roads, and for innumerable other positive misdemeanors. Now these prohibitory laws do not make the transgression a moral offense, or sin: the only obligation in conscience is to submit to the penalty, if levied. It must however be observed, that we are here speaking of laws that are simply and purely penal, where the thing forbidden or enjoined is wholly a matter of indifference, and where the penalty inflicted is an adequate compensation for the civil inconvenience supposed to arise from the offence. But where disobedience to the law involves in it also any degree of public mischief or private injury, there it falls within our former distinction, and is also an offence against conscience.

REVIEW AND DISCUSSION QUESTIONS

1. In the most general sense, what is law, according to Blackstone? What immutable laws of good and evil has God handed down?
2. What is the difference between natural law and divine law (or the law of revelation), and what is their relation to human law?
3. What does Blackstone mean by municipal law, and what are its main characteristics?
4. Blackstone distinguishes between laws forbidding actions that are *mala in se* and those forbidding actions that are *mala prohibita*. Do you agree that we are morally obliged ("bound in conscience") to obey the former, but not the latter?

The Province of Jurisprudence Determined

John Austin

Along with Jeremy Bentham, John Austin was among the most influential of the early utilitarian philosophers. In this selection, taken from his important work on law, Austin explains his positivist conception of law as a species of command and argues that the question of a law's status as law is distinct from the question of whether it is just or compatible with divine law.

LECTURE 1

The matter of jurisprudence is positive law: law, simply and strictly so called: or law set by political superiors to political inferiors. . . .

A law, in the most general and comprehensive acceptation in which the term, in its literal meaning, is employed, may be said to be a rule laid down for the guidance of an intelligent being by an intelligent being having power over him. Under this definition are included, and without impropriety, several species. It is necessary to define accurately the line of demarcation which separates these species from one another, as much mistiness and intricacy has been infused into the science of jurisprudence by their being confounded or not clearly distinguished. In the comprehensive sense above indicated, or in the largest meaning which it has, without extension by metaphor or analogy, the term *law* embraces the following objects:—Laws set by God to his human creatures, and laws set by men to men.

The whole or a portion of the laws set by God to men is frequently styled the law of nature, or natural law: being, in truth, the only natural law of which it is possible to speak without a metaphor, or without a blending of objects which ought to be distinguished broadly. But, rejecting the appellation Law of Nature as ambiguous and misleading, I name those laws or rules, as considered collectively or in a mass, the *Divine law,* or the *law of God.*

Laws set by men to men are of two leading or principal classes: classes which are often blended, although they differ extremely; and which, for that reason, should be severed precisely, and opposed distinctly and conspicuously.

Of the laws or rules set by men to men, some are established by *political* superiors, sovereign and subject: by persons exercising supreme and subordinate *government,* in independent nations, or independent political societies. The aggregate of the rules thus established, or some aggregate forming a portion of that aggregate, is the appropriate matter of jurisprudence, general or particular. To the aggregate of the rules thus established, or to some aggregate forming a portion of that aggregate, the term *law,* as used simply and strictly, is exclusively applied. . . . As contradistinguished to the rules which I style *positive morality,* and so on which I shall touch immediately, the aggregate of the rules, established by political superiors, may also be marked commodiously with the name of *positive law.* . . .

Though *some* of the laws or rules, which are set by men to men, are established by political superiors, *others* are *not* established by political superiors, or are *not* established by political superiors, in that capacity or character.

Closely analogous to human laws of this second class, are a set of objects frequently but *improperly* termed *laws,* being rules set and enforced by *mere opinion,* that is, by the opinions or sentiments held or felt by an indeterminate body of men in regard to human conduct. Instances of such a use of the term *law* are the expressions—"The law of honour"; "The law set by fashion"; and rules of this species constitute much of what is usually termed "International law."

The aggregate of human laws properly so called belonging to the second of the classes above mentioned, with the aggregate of objects *improperly* but by *close analogy* termed laws, I place together in a common class, and denote them by the term *positive morality.* The name *morality* severs them from *positive law,* while the epithet *positive* disjoins them from the *law of God.* And to the end

of obviating confusion, it is necessary or expedient that they *should* be disjoined from the latter by that distinguishing epithet. For the name *morality* (or *morals*), when standing unqualified or alone, denotes indifferently either of the following objects: namely, positive morality *as it is,* or without regard to its merits; and positive morality *as it would be,* if it conformed to the law of God, and were, therefore, deserving of *approbation.*

. . . I shall now state the essentials of *a law* or *rule* (taken with the largest signification which can be given to the term *properly*).

Every *law* or *rule* (taken with the largest signification which can be given to the term *properly*) is a *command.* Or, rather, laws or rules, properly so called, are a *species* of commands.

. . . A command is distinguished from other significations of desire, not by the style in which the desire is signified, but by the power and the purpose of the party commanding to inflict an evil or pain in case the desire be disregarded. If you cannot or will not harm me in case I comply not with your wish, the expression of your wish is not a command, although you utter your wish in imperative phrase. . . .

A command, then, is a signification of desire. But a command is distinguished from other significations of desire by this peculiarity: that the party to whom it is directed is liable to evil from the other, in case he comply not with the desire.

Being liable to evil from you if I comply not with a wish which you signify, I am *bound* or *obliged* by your command, or I lie under a *duty* to obey it. If, in spite of that evil in prospect, I comply not with the wish which you signify, I am said to disobey your command, or to violate the duty which it imposes.

Command and duty are, therefore, correlative terms: the meaning denoted by each being implied or supposed by the other. Or (changing the expression) wherever a duty lies, a command has been signified; and whenever a command is signified, a duty is imposed.

. . . The greater the eventual evil, and the greater the chance of incurring it, the greater is the efficacy of the command, and the greater is the strength of the obligation. . . .

Rewards are, indisputably, *motives* to comply with the wishes of others. But to talk of commands and duties as *sanctioned* or *enforced* by rewards, or to talk of rewards as *obliging* or *constraining* to obedience, is surely a wide departure from the established meaning of the terms. . . .

It appears, then, from what has been premised, that the ideas or notions comprehended by the term *command* are the following. 1. A wish or desire conceived by a rational being, that another rational being shall do or forbear. 2. An evil to proceed from the former, and to be incurred by the latter, in case the latter comply not with the wish. 3. An expression or intimation of the wish by words or other signs.

It also appears from what has been premised, that *command, duty,* and *sanction* are inseparably connected terms: that each embraces the same ideas as the others, though each denotes those ideas in a peculiar order or series. . . .

Now where it obliges *generally* to acts or forbearances of a *class,* a command is a law or rule. But where it obliges to a *specific* act or forbearance, or to acts or forbearances which it determines *specifically* or *individually,* a command is occasional or particular. . . .

If you command your servant to go on a given errand, or *not* to leave your house on a given evening, or to rise at such an hour on such a morning, or to rise at that hour during the next week or month, the command is occasional or particular. For the act or acts enjoined or forbidden are specially determined or assigned.

But if you command him *simply* to rise at that hour, or to rise at that hour *always,* or to rise at that hour *till further orders,* it may be said, with propriety, that you lay down a *rule* for the guidance of your servant's conduct. For no specific act is assigned by the command, but the command obliges him generally to acts of a determined class.

If a regiment be ordered to attack or defend a post, or to quell a riot, or to march from their present quarters, the command is occasional or particular. But an order to exercise daily till further orders shall be given would be called a *general* order, and *might* be called a *rule.*

If Parliament prohibited simply the exportation of corn, either for a given period or indefinitely, it would establish a law or rule: a *kind* or *sort* of act being determined by the command, and acts of that kind or sort being *generally* forbidden. But an order issued by Parliament to meet an impending scarcity, and stopping the exportation

of corn *then shipped and in port,* would not be a law or rule, though issued by the sovereign legislature. . . .

Now the lawgiver determines a class or description of acts; prohibits acts of the class generally and indefinitely; and commands, with the like generality, that punishment shall follow transgression. The command of the lawgiver is, therefore, a law or rule. But the command of the judge is occasional or particular. For he orders a specific punishment, as the consequence of a specific offence. . . .

It appears, from what has been premised, that a law, properly so called, may be defined in the following manner.

A law is a command which obliges a person or persons.

But, as contradistinguished or opposed to an occasional or particular command, a law is a command which obliges a person or persons, and obliges *generally* to acts or forbearances of a *class.*

In language more popular but less distinct and precise, a law is a command which obliges a person or persons to a *course* of conduct.

Laws and other commands are said to proceed from *superiors,* and to bind or oblige *inferiors.* I will, therefore, analyze the meaning of those correlative expressions; and will try to strip them of a certain mystery, by which that simple meaning appears to be obscured.

Superiority is often synonymous with *precedence or excellence.* . . .

But, taken with the meaning wherein I here understand it, the term *superiority* signifies *might:* the power of affecting others with evil or pain, and of forcing them, through fear of that evil, to fashion their conduct to one's wishes.

For example, God is emphatically the *superior* of Man. For His power of affecting us with pain, and of forcing us to comply with His will, is unbounded and resistless.

To a limited extent, the sovereign One or Number is the superior of the subject or citizen: the master, of the slave or servant: the father, of the child.

In short, whoever can *oblige* another to comply with his wishes, is the *superior* of that other, so far as the ability reaches: The party who is obnoxious to the impending evil, being, to that same extent, the *inferior.*

The might or superiority of God, is simple or absolute. . . .

A member of a sovereign assembly is the superior of the judge: the judge being bound by the law which proceeds from that sovereign body. But, in his character of citizen or subject, he is the inferior of the judge: the judge being the minister of the law, and armed with the power of enforcing it.

It appears, then, that the term *superiority* (like the terms *duty* and *sanction*) is implied by the term *command.* For superiority is the power of enforcing compliance with a wish: and the expression or intimation of a wish, with the power and the purpose of enforcing, are the constituent elements of a command.

"That *laws* emanate from *superiors*" is, therefore, an identical proposition. For the meaning which it affects to impart is contained in its subject. . . .

According to an opinion which I must notice *incidentally* here, though the subject to which it relates will be treated *directly* hereafter, *customary laws* must be excepted from the proposition "that laws are a series of commands."

By many of the admirers of customary laws . . . they are thought to oblige legally (independently of the sovereign or state), *because* the citizens or subjects have observed or kept them. . . .

At its origin, a custom is a rule of conduct which the governed observe spontaneously, or not in pursuance of a law set by a political superior. The custom is transmitted into positive law, when it is adopted as such by the courts of justice, and when the judicial decisions fashioned upon it are enforced by the power of the state. But before it is adopted by the courts, and clothed with the legal sanction, it is merely a rule of positive morality: a rule generally observed by the citizens or subjects; but deriving the only force, which it can be said to possess, from the general disapprobation falling on those who transgress it.

Now when judges transmute a custom into a legal rule (or make a legal rule not suggested by a custom), the legal rule which they establish is established by the sovereign legislature. A subordinate or subject judge is merely a minister. The portion of the sovereign power which lies at his disposition is merely delegated. The rules which he makes derive their legal force from authority given by the state: an authority which the state may confer

expressly, but which it commonly imparts in the way of acquiescence. For, since the state may reverse the rules which he makes, and yet permits him to enforce them by the power of the political community, its sovereign will "that his rules shall obtain as law" is clearly evinced by its conduct, though not by its express declarations. . . .

LECTURE 5

. . . Now it follows from these premises, that the laws of God, and positive laws are laws proper, or laws properly so called.

The laws of God are laws proper, inasmuch as they are *commands* express or tacit, and therefore emanate from a *certain* source.

Positive laws, or laws strictly so called, are established directly or immediately by authors of three kinds:—by monarchs, or sovereign bodies, as supreme political superiors: by men in a state of subjection, as subordinate political superiors: by subjects as private persons, in pursuance of legal rights. But every positive law, or every law strictly so called, is a direct or circuitous command of a monarch or sovereign number in the character of political superior: that is to say, a direct or circuitous command of a monarch or sovereign number to a person or persons in a state of subjection to its author. And being a *command* (and therefore flowing from a *determinate* source), every positive law is a law proper, or a law properly so called.

Besides the human laws which I style positive law, there are human laws which I style positive morality, rules of positive morality, or positive moral rules.

The generic character of laws of the class may be stated briefly in the following negative manner—No law belonging to the class is a direct or circuitous command of a monarch or sovereign number in the character of political superior. In other words, no law belonging to the class is a direct or circuitous command of a monarch or sovereign number to a person or persons in a state of subjection to its author. . . .

The existence of law is one thing; its merit or demerit is another. Whether it be or be not is one enquiry; whether it be or be not conformable to an assumed standard, is a different enquiry. A law, which actually exists, is a law, though we happen to dislike it, or though it vary from the text, by which we regulate our approbation and disapprobation. This truth, when formally announced as an abstract proposition, is so simple and glaring that it seems idle to insist upon it. But simple and glaring as it is, when enunciated in abstract expressions the enumeration of the instances in which it has been forgotten would fill a volume.

Sir William Blackstone, for example, says in his "Commentaries," that the laws of God are superior in obligation to all other laws; that no human laws should be suffered to contradict them; that human laws are of no validity if contrary to them; and that all valid laws derive their force from that Divine original.

Now, he *may* mean that all human laws ought to conform to the Divine laws. If this be his meaning, I assent to it without hesitation. The evils which we are exposed to suffer from the hands of God as a consequence of disobeying His commands are the greatest evils to which we are obnoxious; the obligations which they impose are consequently paramount to those imposed by any other laws, and if human commands conflict with the Divine law, we ought to disobey the command which is enforced by the less powerful sanction; this is implied in the term *ought:* the proposition is identical, and therefore perfectly indisputable—it is our interest to choose the smaller and more uncertain evil, in preference to the greater and surer. If this be Blackstone's meaning, I assent to his proposition, and have only to object to it, that it tells us just nothing.

Perhaps, again, he means that human lawgivers are themselves obliged by the Divine laws to fashion the laws which they impose by that ultimate standard, because if they do not, God will punish them. To this also I entirely assent. . . .

But the meaning of this passage of Blackstone, if it has a meaning, seems rather to be this: that no human law which conflicts with the Divine law is obligatory or binding; in other words, that no human law which conflicts with the Divine law is a *law,* for a law without an obligation is a contradiction in terms. I suppose this to be his meaning, because when we say of any transaction that it is invalid or void, we mean that it is not binding: as, for example, if it be a contract, we mean that the political law will not lend its sanction to enforce the contract.

Now, to say that human laws which conflict with the Divine law are not binding, that is to say,

are not laws, is to talk stark nonsense. The most pernicious laws, and therefore those which are most opposed to the will of God, have been and are continually enforced as laws by judicial tribunals. Suppose an act innocuous, or positively beneficial, be prohibited by the sovereign under the penalty of death; if I commit this act, I shall be tried and condemned, and if I object to the sentence, that it is contrary to the law of God, who has commanded that human lawgivers shall not prohibit acts which have no evil consequences, the Court of Justice will demonstrate the inconclusiveness of my reasoning by hanging me up, in pursuance of the law of which I have impugned the validity. An exception, demurrer, or plea, founded on the law of God was never heard in a Court of Justice, from the creation of the world down to the present moment.

But this abuse of language is not merely puerile, it is mischievous. When it is said that a law ought to be disobeyed, what is meant is that we are urged to disobey it by motives more cogent and compulsory than those by which it is itself sanctioned. If the laws of God are certain, the motives which they hold out to disobey any human command which is at variance with them are paramount to all others. But the laws of God are not always certain. . . . In quiet times the dictates of utility are fortunately so obvious that the anarchical doctrine sleeps, and men habitually admit the validity of laws which they dislike. To prove by pertinent reasons that a law is pernicious is highly useful, because such process may lead to the abrogation of the pernicious law. To incite the public to resistance by determinate views of *utility* may be useful, for resistance, grounded on clear and definite prospects of good, is sometimes beneficial. But to proclaim generally that all laws which are pernicious or contrary to the will of God are void and not to be tolerated, is to preach anarchy, hostile and perilous as much to wise and benign rule as to stupid and galling tyranny. . . .

LECTURE 6

. . . Every positive law, or every law simply and strictly so called, is set by a sovereign person, or a sovereign body of persons, to a member or members of the independent political society wherein that person or body is sovereign or supreme. Or (changing the expression) it *is* set by a monarch, or sovereign number, to a person or persons in a state of subjection to its author. Even though it sprung directly from another fountain or source, it is a positive law, or a law strictly so called, by the institution of that present sovereign in the character of political superior. . . .

The superiority which is styled sovereignty, and the independent political society which sovereignty implies, is distinguished from other superiority, and from other society, by the following marks or characters—1. The bulk of the given society are in a *habit* of obedience or submission to a *determinate* and *common* superior: let that common superior be a certain individual person, or a certain body or aggregate of individual persons. 2. That certain individual, or that certain body of individuals, is *not* in a habit of obedience to a determinate human superior. Laws (improperly so called) which opinion sets or imposes, may permanently affect the conduct of that certain individual or body. To express or tacit commands of other determinate parties, that certain individual or body may yield occasional submission. But there is no determinate person, determinate aggregate of persons, to whose commands, express or tacit, that certain individual or body renders habitual obedience. . . . By "an independent political society," or "an independent and sovereign nation," we mean a political society consisting of a sovereign and subjects, as opposed to a political society which is merely subordinate: that is to say, which is merely a limb or member of another political society, and which therefore consists entirely of persons in a state of subjection.

In order that a given society may form a society political and independent, the two distinguishing marks which I have mentioned above must unite. The *generality* of the given society must be in the *habit* of obedience to a *determinate* and *common* superior: whilst that determinate person, or determinate body of persons must *not* be habitually obedient to a determinate person or body. It is the union of that positive, with this negative mark, which renders that certain superior sovereign or supreme, and which renders that given society (including that certain superior) a society political and independent.

REVIEW AND DISCUSSION QUESTIONS

1. Explain the two types of human law. (The second kind also includes "laws improperly so-called.")
2. Laws are rules, according to Austin, but rules are really just a species of command. How does Austin define a command?
3. Not all commands are laws. Explain.
4. How does Austin understand the notion of duty or obligation?
5. Which is superior, legal obligation or religious obligation? Explain.
6. Austin thinks that international law is "improperly" termed "law." Why do you think he believes this?
7. Austin thinks Blackstone's view is dangerous ("mischievous"). Explain.
8. How does Austin account for common or judge-made law?
9. Suppose that a judge, noticing there is no relevant statute, applies a customary rule that people have followed for a generation and holds a ship liable for not coming to the aid of another at sea. Who is the sovereign in that case? Was the customary rule law before the judge applied it according to Austin?
10. Every positive law is set by the sovereign (which can be a person or a body of persons). What makes something the "sovereign"?
11. Elsewhere Austin argues that in a democracy the electorate is the sovereign. Do you agree? If the people are the sovereign, does it still make sense to think of laws as commands?
12. Austin's most famous statement, "The existence of law is one thing; its merit or demerit is another," is often seen as defining legal positivism. How do you interpret it?

CHAPTER 8

Formalism and Legal Realism

Judges and lawyers often disagree about how a statute or precedent should be interpreted. Such disagreements raise important questions. Is there really a right or best answer to legal questions, or are there instead only different, subjective opinions? Are past precedents, statutes, and other legal materials sufficient to determine the correct legal outcome in a given case? Or is it instead merely a matter of a particular judge's personal values or political opinions? These important questions engage the authors presented in this section, but they will also emerge in various other contexts throughout the book. The term "formalist" is often used to describe those who think that if properly interpreted, the law speaks with a single voice, so there is in theory always a right legal answer. This formalist position is represented by Joseph Beale. Opposing the formalist view is a school of thinking known as "legal realism." Legal realists deny that the law is determinate. In the early decades of the twentieth century, legal realism emerged as a distinctive and radical school of jurisprudence in the United States, and today it continues to exert an important influence on legal philosophy and legal scholarship. Legal realists reacted against what they saw as the "mechanical" and "nonpolitical" approach of those who saw the law as a logical and consistent system of rules and principles. Realists such as Oliver Wendell Holmes and Jerome Frank thus adopted a more pragmatic approach to law, an approach frequently informed by the social sciences and eager to acknowledge the extent to which various social forces external to the law influence its development. Realists see the law as consisting of decisions, not rules. Judges do not apply the law. They make it.

The System of Law

Joseph H. Beale

In this brief selection from *A Treatise on the Conflict of Laws,* which discusses how the law addresses conflicts among its various compartments and jurisdictions, Joseph Beale summarizes some of the central features of the formalist's picture of law. Properly understood, law is a system or body that includes

various sorts of standards, rules, and principles of conduct. Taken together, these constitute a "single, homogeneous philosophical system." While law and legal principles can be changed either by legislation or by bad judicial decisions, Beale argues that these are not pervasive enough to render the entire system incoherent but instead represent what is "local" about each system. Beale concludes with a discussion of the general characteristics of a legal system: generality, universality, continuity, and predictability.

Body of Principles, Standards, and Rules. The particular law of a state is made up of several elements, some of them quite arbitrary in their operation. These accidental elements prevent the law from becoming, like the common law a scientific system.

The principles which form the largest portion even of the particular law of a state are general premises of law which can be used for deduction and for analogy. Such principles are: "A conveyance of any interest in land is governed as to its validity by the law of the state of situs." "The existence of a cause of action for a tort is governed by the law of the place of wrong."

A standard is a rule which is stated as a degree of a continuously changing series to be reached, in order for a legal result to follow; the application of the standard requires therefore not merely the finding of facts but the finding of this peculiar fact, the reaching of a particular degree in events which are capable of a continuous series. Examples of standards are: a negligence is the lack of *due care;* a result of an act of which there was *appreciable risk* is proximate.

A rule means a statement of law applicable only to a narrowly defined class of cases and incapable of extension by deduction or analogy. Instances of rules are: the rule in Shelley's case; the rule that one must stop, look, and listen before crossing a railroad track; the rule that one must turn to the right to pass a vehicle coming in an opposite direction. Most statutory law is of this sort. . . .

Definition of Law. The writers of the analytical school of jurisprudence, emphasizing the positive character of law as an expression of sovereign will, have proposed definitions which fit one portion of the law only; that is, the rules made by the legislative body. . . .

These definitions appear to ignore the principal element of law, the so-called "unwritten law." So important is this portion of the law, and so widespread is the distinction, that in other languages than English different terms are used to distinguish the two. The positive law formulated and fixed by a legislative body is called *lex, loi, Gesetz;* the general unwritten law is called *ius, droit, Recht.*

The definitions of the analytic school are properly applicable only to *lex.* Austin, realizing this defect, fitted the facts to his theory by assuming a tacit command by the sovereign to his judges to express the rules of law which they lay down in their decisions; thus assimilating judicial to statute law.

A second objection to these definitions, even extended (by a fiction) to cover the unwritten law, is that they all ignore that quality of the law which is absolutely characteristic: that it tends to form a single homogeneous philosophical system. Any definition of law which treats each part of it as an isolated thing, instead of as part of an embodied system, misses its nature altogether. . . .

The authors whose opinions have been examined neglect, or at least too little emphasize, the one most important feature of law: that it is not a mere collection of arbitrary rules, but a body of scientific principle. That part of the law which in other languages is known as *ius, Recht, droit,* is a branch of practical philosophy; by which, through the use of reason and experience, legal generalizations may be made. Purity of doctrine may be lost through wrong decisions of courts, thus warping legal principle by bad precedent; but wrong decisions are after all uncommon, and the law is not seriously affected by them. The application of general principles may be inhibited by legislation: but the amount of legislation which affects ordinary private law is relatively small, and doctrine is not greatly changed by statute. Much the largest and most important part of the law, therefore, is this body of principle, or as it is almost invariably called by European writers, *doctrine.* The changes in principle made by legislation and by wrong decisions

constitute the greater part of the peculiar local law of any jurisdiction, as distinguished from the general doctrine of the prevailing legal system.

Law, therefore, is made in part by the legislature; in part it rests upon precedent; and in great part it consists in a homogeneous, scientific, and all-embracing body of principle; and a correct definition of law in general must apply to all these varieties of law. . . .

The Characteristics of Law. An essential characteristic of law is its generality; since justice requires equality of treatment for all persons, and this means generality. It is, as has been seen, a body of general principles, not a collection of special commands. A set of rules for the action of a particular person would lack this character of generality. Thus, the decision and judgment of a court, determining a particular controversy and laying an order upon one party to it, lacks this element and can in no sense be regarded as in itself law. . . . Law being a general principle applying indifferently to all cases which in the future can arise under it, the decision of a court can be law only if the court has power in its decision to lay down binding rules for future conduct. . . .

Another characteristic of law is universality. It is unthinkable in a civilized country that any act should fall outside of the domain of law. If law be regarded as a command, then every act done must either be permitted or forbidden. . . .

Another characteristic of law is continuity. "From the day of its promulgation to the day of its repeal the law must always be heard and obeyed."[1]

There can be no break or interregnum in law. From the time law comes into existence with the first felt corporateness of a primitive people, it must last until the final disappearance of human society. Once created, it persists until a change takes place, and when changed it continues in such changed condition until the next change, and so on forever. . . .

The social need of continuity in law is most clearly felt because society needs to know the law in advance of judicial action upon it. In order that law may help rather than hinder the carrying on of the work of society it must be possible for every person, of his own knowledge or by the help of others' knowledge, to discover the application of the law to any contemplated act. He must be in some way secured against unexpected legal consequences of his actions. Business could not go on, industry could not be maintained unless it were possible for the producer or the merchant to learn how he could conform his activities to the law. For this purpose it must be possible for one learned in the law to speak with authority on the application of law to the proposed acts and to predict with reasonable degree of certainty the decision of courts in case the legality of the acts should be called in question. If there were any discontinuity in the law,—if, for instance, a judicial tribunal had the power to change the law as it liked, or the discretion as to the application of law to the facts,—the client would seek advice in vain, for counsel however learned could only vaguely guess what the law would be at the time of possible future litigation. . . .

In order that it may be possible, then, for the law to function, it should be predicable; that is, in a case not previously passed on by a court, competent counsel should be able to predict what the action of a court in the case will be. Predicability can be secured only by a union of regularity, system, and reason. [Beale does not include a separate discussion of "reason" apart from "regularity" and "system"—EDS.]

a. Regularity. An administrative tribunal may guess at the "very right" of the particular case before it; but since each man's reaction to the idea of right is subjective to himself, it is practically impossible in a given case to predict the reaction of some as yet undetermined body of men to a given state of fact. To be predicable, therefore, the result of judicial action must not lie in mere judicial discretion. That dictum of the fourteenth-century judge, that law is the will of the judges, his more experienced colleague at once corrected by pointing out a criterion of the decision: "By no means; it is the rule of right."[2] The court must have a rule to guide it, or its judgment, being merely discretionary, will be impredicable.

b. System. It is not enough, however, that law should be regular; for a merely regular law fails to meet the needs of modern civilization. A mass of rules is necessarily limited in its application; no legislator can formulate rules which will cover all future cases. Law must have the power of extension to novel facts. An example of this need

occurred when streetcars were first introduced. A street-car line between Charlestown and Boston was opened about 1856, at a time when one man enjoyed a monopoly of public conveyance between the cities by omnibus and dray. The drays thereupon drove slowly along the tracks of the street-cars, while the busses hurried by. No formulated rule applied to the case, for it could not have been foreseen and guarded against. If the law had been regular but non-systematic, like the folk-law that preceded the common law, the wrong lacking a formal remedy, must have been unredressed. But the common law was systematic; that is, it consisted of a system of thought based upon principles which covered every possible occurrence. Every human act was either permitted or forbidden; every act either changed or left unchanged existing rights. Under this system of rights the act of a drayman in delaying the street-car was either forbidden or permitted by the law, whether the rule that determined the answer had ever been formulated or not; and the court was able to say that legal principle forbade the act of the drayman.

NOTES

1. *Commonwealth v. Temple,* 80 Mass. (14 Gray) 69 (1860) (Shaw, C. J.).
2. Ibid.

REVIEW AND DISCUSSION QUESTIONS

1. Beale claims that law includes rules and principles. How does he distinguish the two?
2. Why does Beale think that those who see law as the expression of rules of the legislative body or of the sovereign will have too narrow a view of the law?
3. Explain the importance of the unwritten law, for Beale.
4. Describe the essential characteristics of the law, according to Beale. Why does he call it a "system"?

The Path of the Law

Oliver Wendell Holmes

Written in 1897, this much-discussed essay has seemed to many to capture the central tenets of what came to be called legal realism. In it, Oliver Wendell Holmes offers a perspective on law at variance with both the formalists who had gone before and the traditional positivists, such as John Austin. According to Holmes, the law is, in fact, nothing more than past decisions plus predictions of what future judges will do. In making that claim, he sets himself against not only those who see law as a coherent, complete system of rules and principles, but also against people who seek to understand the law as commands or orders of a sovereign power.

When we study law we are not studying a mystery but a well known profession. We are studying what we shall want in order to appear before judges, or to advise people in such a way as to keep them out of court. The reason why it is a profession, why people will pay lawyers to argue for them or to advise them is that in societies like ours the command of the public force is entrusted to the judges in certain cases, and the whole power of the state will be put forth, if necessary, to carry out their

judgments and decrees. People want to know under what circumstances and how far they will run the risk of coming against what is so much stronger than themselves, and hence it becomes a business to find out when this danger is to be feared. The object of our study, then, is prediction, the prediction of the incidence of the public force through the instrumentality of the courts.

The means of the study are a body of reports, of treatises, and of statutes, in this country and in England, extending back for six hundred years, and now increasing annually by hundreds. In these sibylline leaves are gathered the scattered prophecies of the past upon the cases in which the axe will fall. These are what properly have been called the oracles of the law. Far the most important and pretty nearly the whole meaning of every new effort of legal thought is to make these prophecies more precise, and to generalize them into a thoroughly connected system. The process is one, from a lawyer's statement of a case, eliminating as it does all the dramatic elements with which his client's story has clothed it, and retaining only the facts of legal import, up to the final analyses and abstract universals of theoretic jurisprudence. The reason why a lawyer does not mention that his client wore a white hat when he made a contract, while Mrs. Quickly would be sure to dwell upon it along with the parcel gilt goblet and the sea-coal fire is that he forsees that the public force will act in the same way whatever his client had upon his head. It is to make the prophecies easier to be remembered and to be understood that the teachings of the decisions of the past are put into general propositions and gathered into text-books, or that statutes are passed in a general form. . . .

The number of our predictions when generalized and reduced to a system is not unmanageably large. They present themselves as a finite body of dogma which may be mastered within a reasonable time. . . .

The first thing for a business-like understanding of the matter is to understand its limits, and therefore I think it desirable at once to point out and dispel a confusion between morality and law, which sometimes rises to the height of conscious theory, and more often and indeed constantly is making trouble in detail without reaching the point of consciousness. You can see very plainly that a bad man has as much reason as a good one for wishing to avoid an encounter with the public force, and therefore you can see the practical importance of the distinction between morality and law. A man who cares nothing for an ethical rule which is believed and practised by his neighbors is likely nevertheless to care a good deal to avoid being made to pay money, and will want to keep out of jail if he can.

I take it for granted that no hearer of mine will misinterpret what I have to say as the language of cynicism. The law is the witness and external deposit of our moral life. Its history is the history of the moral development of the race. The practice of it, in spite of popular jests, tends to make good citizens and good men. When I emphasize the difference between law and morals I do so with reference to a single end, that of learning and understanding the law. For that purpose you must definitely master its specific marks, and it is for that that I ask you for the moment to imagine yourselves indifferent to other and greater things. . . .

If you want to know the law and nothing else, you must look at it as a bad man, who cares only for the material consequences which such knowledge enables him to predict, not as a good one, who finds his reasons for conduct, whether inside the law or outside of it, in the vaguer sanctions of conscience. The theoretical importance of the distinction is no less, if you would reason on your subject aright. The law is full of phraseology drawn from morals, and by the mere force of language continually invites us to pass from one domain to the other without perceiving it, as we are sure to do unless we have the boundary constantly before our minds. The law talks about rights, and duties, and malice, and intent, and negligence, and so forth, and nothing is easier, or, I may say, more common in legal reasoning, than to take these words in their moral sense, at some stage of the argument, and so to drop into fallacy. For instance, when we speak of the rights of man in a moral sense, we mean to mark the limits of interference with individual freedom which we think are prescribed by conscience, or by our ideal, however reached. Yet it is certain that many laws have been enforced in the past, and it is likely that some are enforced now, which are condemned by the most enlightened opinion of the time, or which at all events pass the limit of interference as many consciences would draw it.

Manifestly, therefore, nothing but confusion of thought can result from assuming that the rights of man in a moral sense are equally rights in the sense of the Constitution and the law. No doubt simple and extreme cases can be put of imaginable laws which the statute-making power would not dare to enact, even in the absence of written constitutional prohibitions, because the community would rise in rebellion and fight; and this gives some plausibility to the proposition that the law, if not a part of morality, is limited by it. But this limit of power is not coextensive with any system of morals. For the most part it falls far within the lines of any such system, and in some cases may extend beyond them, for reasons drawn from the habits of a particular people at a particular time. I once heard the late Professor Agassiz say that a German population would rise [in rebellion] if you added two cents to the price of a glass of beer. A statute in such a case would be empty words, not because it was wrong, but because it could not be enforced. No one will deny that wrong statutes can be and are enforced, and we should not all agree as to which were the wrong ones.

The confusion with which I am dealing besets confessedly legal conceptions. Take the fundamental question, What constitutes the law? You will find some text writers telling you that it is something different from what is decided by the courts of Massachusetts or England, that it is a system of reason, that it is a deduction from principles of ethics or admitted axioms or what not, which may or may not coincide with the decisions. But if we take the view of our friend the bad man we shall find that he does not care two straws for the axioms or deductions, but that he does want to know what the Massachusetts or English courts are likely to do in fact. I am much of his mind. The prophecies of what the courts will do in fact, and nothing more pretentious, are what I mean by the law.

Take again a notion which as popularly understood is the widest conception which the law contains;—the notion of legal duty, to which already I have referred. We fill the word with all the content which we draw from morals. But what does it mean to a bad man? Mainly, and in the first place, a prophecy that if he does certain things he will be subjected to disagreeable consequences by way of imprisonment or compulsory payment of money. But from his point of view, what is the difference between being fined and being taxed a certain sum for doing a certain thing? That his point of view is the test of legal principles is shown by the many discussions which have arisen in the courts on the very question whether a given statutory liability is a penalty or a tax. On the answer to this question depends the decision whether conduct is legally wrong or right, and also whether a man is under compulsion or free. Leaving the criminal law on one side, what is the difference between the liability under the mill acts or statutes authorizing a taking by eminent domain and the liability for what we call a wrongful conversion of property where restoration is out of the question? In both cases the party taking another man's property has to pay its fair value as assessed by a jury, and no more. What significance is there in calling one taking right and another wrong from the point of view of the law? It does not matter, so far as the given consequence, the compulsory payment, is concerned, whether the act to which it is attached is described in terms of praise or in terms of blame, or whether the law purports to prohibit it or allow it. . . .

Nowhere is the confusion between legal and moral ideas more manifest than in the law of contract. Among other things, here again the so called primary rights and duties are invested with a mystic significance beyond what can be assigned and explained. The duty to keep a contract at common law means a prediction that you must pay damages if you do not keep it,—and nothing else. If you commit a tort, you are liable to pay a compensatory sum. If you commit a contract, you are liable to pay a compensatory sum unless the promised event comes to pass, and that is all the difference. But such a mode of looking at the matter stinks in the nostrils of those who think it advantageous to get as much ethics into the law as they can. . . .

We talk about a contract as a meeting of the minds of the parties, and thence it is inferred in various cases that there is no contract because their minds have not met; that is, because they have intended different things or because one party has not known of the assent of the other. Yet nothing is more certain than that parties may be bound by a contract to things which neither of them intended, and when one does not know of the other's assent. Suppose a contract is executed in due form and in

writing to deliver a lecture, mentioning no time. One of the parties thinks that the promise will be construed to mean at once, within a week. The other thinks that it means when he is ready. The court says that it means within a reasonable time. The parties are bound by the contract as it is interpreted by the court, yet neither of them meant what the court declares that they have said. . . .

I hope that my illustrations have shown the danger, both to speculation and to practice, of confounding morality with law, and the trap which legal language lays for us on that side of our way. For my own part, I often doubt it would not be a gain if every word of moral significance could be banished from the law altogether, and other words adopted which should convey legal ideas uncolored by anything outside the law. We should lose the fossil records of a good deal of history and the majesty got from ethical associations, but by ridding ourselves of an unnecessary confusion we should gain very much in the clearness of our thought.

So much for the limits of the law. The next thing which I wish to consider is what are the forces which determine its content and its growth. You may assume, with Hobbes and Bentham and Austin, that all law emanates from the sovereign, even when the first human beings to enunciate it are the judges, or you may think that law is the voice of the Zeitgeist, or what you like. It is all one to my present purpose. Even if every decision required the sanction of an emperor with despotic power and a whimsical turn of mind, we should be interested none the less, still with a view to prediction, in discovering some order, some rational explanation, and some principle of growth for the rules which he laid down. In every system there are such explanations and principles to be found. It is with regard to them that a second fallacy comes in, which I think it important to expose.

The fallacy to which I refer is the notion that the only force at work in the development of the law is logic. . . . The danger of which I speak is not the admission that the principles governing other phenomena also govern the law, but the notion that a given system, ours, for instance, can be worked out like mathematics from some general axioms of conduct. This is the natural error of the schools, but it is not confined to them. I once heard a very eminent judge say that he never let a decision go until he was absolutely sure that it was right. So judicial dissent often is blamed, as if it meant simply that one side or the other were not doing their sums right, and, if they would take more trouble, agreement inevitably would come.

This mode of thinking is entirely natural. The training of lawyers is a training in logic. The processes of analogy, discrimination, and deduction are those in which they are most at home. The language of judicial decision is mainly the language of logic. And the logical method and form flatter that longing for certainty and for repose which is in every human mind. But certainty generally is illusion, and repose is not the destiny of man. Behind the logical form lies a judgment as to the relative worth and importance of competing legislative grounds, often an inarticulate and unconscious judgment, it is true, and yet the very root and nerve of the whole proceeding. You can give any conclusion a logical form. You always can imply a condition in a contract. But why do you imply it? It is because of some belief as to the practice of the community or of a class, or because of some opinion as to policy, or, in short, because of some attitude of yours upon a matter not capable of exact quantitative measurement, and therefore not capable of founding exact logical conclusions. Such matters really are battle grounds where the means do not exist for determinations that shall be good for all time, and where the decision can do no more than embody the preference of a given body in a given time and place. We do not realize how large a part of our law is open to reconsideration upon a slight change in the habit of the public mind. No concrete proposition is self-evident, no matter how ready we may be to accept it, not even Mr. Herbert Spencer's "Every man has a right to do what he wills, provided he interferes not with a like right on the part of his neighbors."

Why is a false and injurious statement privileged, if it is made honestly in giving information about a servant? It is because it has been thought more important that information should be given freely, than that a man should be protected from what under other circumstances would be an actionable wrong. Why is a man at liberty to set up a business which he knows will ruin his neighbor? It is because the public good is supposed to be best subserved by free competition. Obviously such judgments of relative importance may vary in different times and places. . . .

REVIEW AND DISCUSSION QUESTIONS

1. Explain how Holmes understands the essential nature of the law. What, finally, is law, according to him?
2. Why does Holmes suggest that the best way to look at the law is from the perspective of the bad man?
3. How does Holmes explain the distinction between moral and legal duty?
4. What value or values do you think underlie the law, according to Holmes?
5. How do you think Holmes would respond to a legal formalist such as Beale?

Realism and the Law

Jerome Frank

Legal realism arose as a distinctive school of American jurisprudence in the early decades of the twentieth century. Describing their opponents as "formalists," legal realists argued that the law is essentially normative and that judges exercise wide discretion; a judge's values and beliefs about society have a large impact, they argued, on how cases are decided. Law is not a "seamless web" and, more important, it places few rational restraints on how judges rule. Jerome Frank, who served as an appeals court judge, was among the most outspoken of the legal realists. In the following essay, taken from his book *Law and the Modern Mind,* he first describes the "conventional" view of law, which fails to acknowledge that judges often make new law. He then describes how the judge's personality affects the laws he or she makes, and attacks the "myth" that precedents and reason control judges. He concludes with a brief discussion of the nature of law from the perspective of the legal realist.

JUDICIAL LAW-MAKING

Have judges the right and power to make law and change law? Much good ink has been spilled in arguing that question. A brief survey of the controversy will illuminate our thesis.

The conventional view may be summarized thus:

Law is a complete body of rules existing from time immemorial and unchangeable except to the limited extent that legislatures have changed the rules by enacted statutes. Legislatures are expressly empowered thus to change the law. But the judges are not to make or change the law but to apply it. The law, ready-made, pre-exists the judicial decisions.

Judges are simply "living oracles" of law. They are merely "the speaking law." Their function is purely passive. They are "but the mouth which pronounces the law." They no more make or invent new law than Columbus made or invented America. Judicial opinions are evidence of what the law is; the best evidence, but no more than that. When a former decision is overruled, we must not say that the rule announced in the earlier decision was once the law and has now been changed by the later decision. We must view the earlier decision as laying down an erroneous rule. It was a false map of the law just as a pre-Columbian map of the world was false. Emphatically, we must not refer to the new decision as making new law. It only seems to do so. It is merely a bit of revised legal cartography.

If a judge actually attempted to contrive a new rule, he would be guilty of usurpation of power, for the legislature alone has the authority to change the law. The judges, writes Blackstone, are "not delegated to pronounce a new law, but to maintain and expound the old law"; even when a former decision is abandoned because "most evidently contrary to reason," the "subsequent judges do not

pretend to make new law, but to vindicate the old one from misrepresentation." The prior judge's eyesight had been defective and he made "a mistake" in finding the law, which mistake is now being rectified by his successors.

Such is the conventional notion. There is a contrary minority view, which any dispassionate observer must accept as obviously the correct view:

"No intelligent lawyer would in this day pretend that the decisions of the courts do not add to and alter the law," says Pollock, a distinguished English jurist. "Judge-made law is real law," writes Dicey, another famous legal commentator, "though made under the form of, and often described by judges no less than jurists, as the mere interpretation of law. . . . The amount of such judge-made law is in England far more extensive than a student realizes. Nine-tenths, at least, of the law of contract, and the whole, or nearly the whole, of the law of torts are not to be discovered in any volume of the statutes. . . . Whole branches, not of ancient but of very modern law, have been built up, developed or created by action of the courts."

Judges, then, do make and change law. The minority view is patently correct; the opposing arguments will not bear analysis. What, then, explains the belief so tenaciously held that the judiciary does not ever change the law or that, when it does, it is acting improperly? Why is it that judges adhere to what Morris Cohen has happily called "the phonographic theory of the judicial function"? What explains the recent remark of an eminent member of the Bar: "The man who claims that under our system courts make law is asserting that the courts habitually act unconstitutionally"? Why do the courts customarily deny that they have any law-making power and describe new law which they create to deal with essentially contemporary events, as mere explanations or interpretations of law which already exists and has existed from time immemorial? Why this obstinate denial of the juristic realities?

We revert to our thesis: The essence of the basic legal myth or illusion is that law can be entirely predictable. Back of this illusion is the childish desire to have a fixed father-controlled universe, free of chance and error due to human fallibility. . . .

But if it is once recognized that a judge, in the course of deciding a case, can for the first time create the law applicable to that case, or can alter the rules which were supposed to exist before the case was decided, then it will also have to be recognized that the rights and obligations of the parties to that case may be decided retroactively. A change thus made by a judge, when passing upon a case, is a change in the law made with respect to past events,—events which occurred before the law came into existence. Legal predictability is plainly impossible, if, at the time I do an act, I do so with reference to law which, should a lawsuit thereafter arise with reference to my act, may be changed by the judge who tries the case. For then the result is that my case is decided according to law which was not in existence when I acted and which I, therefore, could not have known, predicted or relied on when I acted.

If, therefore, one has a powerful need to believe in the possibility of anything like exact legal predictability, he will find judicial lawmaking intolerable and seek to deny its existence.

Hence the myth that the judges have no power to change existing law or make new law. . . .

THE JUDGING PROCESS AND THE JUDGE'S PERSONALITY

. . . As the word indicates, the judge in reaching a decision is making a judgment. And if we would understand what goes into the creating of that judgment, we must observe how ordinary men dealing with ordinary affairs arrive at their judgments.

The process of judging, so the psychologists tell us, seldom begins with a premise from which a conclusion is subsequently worked out. Judging begins rather the other way around—with a conclusion more or less vaguely formed; a man ordinarily starts with such a conclusion and afterwards tries to find premises which will substantiate it. If he cannot, to his satisfaction, find proper arguments to link up his conclusion with premises which he finds acceptable, he will, unless he is arbitrary or mad, reject the conclusion and seek another.

In the case of the lawyer who is to present a case to a court, the dominance in his thinking of the conclusion over the premises is moderately obvious. He is a partisan working on behalf of his client. The conclusion is, therefore, not a matter of

choice except within narrow limits. He must, that is if he is to be successful, begin with a conclusion which will insure his client's winning the lawsuit. He then assembles the facts in such a fashion that he can work back from this result he desires to some major premise which he thinks the court will be willing to accept. The precedents, rules, principles and standards to which he will call the court's attention constitute this premise.

While "the dominance of the conclusion" in the case of the lawyer is clear, it is less so in the case of the judge. For the respectable and traditional descriptions of the judicial judging process admit no such backward-working explanation. In theory, the judge begins with some rule or principle of law as his premise, applies this premise to the facts, and thus arrives at his decision.

Now, since the judge is a human being and since no human being in his normal thinking process arrives at decisions (except in dealing with a limited number of simple situations) by the route of any such syllogistic reasoning, it is fair to assume that the judge, merely by putting on the judicial ermine, will not acquire so artificial a method of reasoning. Judicial judgments, like other judgments, doubtless, in most cases, are worked out backward from conclusions tentatively formulated.

As Jastrow says, "In spite of the fact that the answer in the book happens to be wrong, a considerable portion of the class succeeds in reaching it. . . . The young mathematician will manage to obtain the answer which the book requires, even at the cost of a resort to very unmathematical processes." Courts, in their reasoning, are often singularly like Jastrow's young mathematician. Professor Tulin has made a study which prettily illustrates that fact. While driving at a reckless rate of speed, a man runs over another, causing severe injuries. The driver of the car is drunk at the time. He is indicted for the statutory crime of "assault with intent to kill." The question arises whether his act constitutes that crime or merely the lesser statutory crime of "reckless driving." The courts of several states have held one way, and the courts of several other states have held the other.

The first group maintain that a conviction for assault with intent to kill cannot be sustained in the absence of proof of an actual purpose to inflict death. In the second group of states the courts have said that it was sufficient to constitute such a crime if there was a reckless disregard of the lives of others, such recklessness being said to be the equivalent of actual intent.

With what, then, appears to be the same facts before them, these two groups of courts seem to have sharply divided in their reasoning and in the conclusions at which they have arrived. But upon closer examination it has been revealed by Tulin that, in actual effect, the results arrived at in all these states have been more or less the same. In Georgia, which may be taken as representative of the second group of states, the penalty provided by the statute for reckless driving is far less than that provided, for instance, in Iowa, which is in the first group of states. If, then, a man is indicted in Georgia for reckless driving while drunk, the courts can impose on him only a mild penalty; whereas in Iowa the judge, under an identically worded indictment, can give a stiff sentence. In order to make it possible for the Georgia courts to give a reckless driver virtually the same punishment for the same offense as can be given by an Iowa judge, it is necessary in Georgia to construe the statutory crime of assault with intent to kill so that it will include reckless driving while drunk; if, and only if, the Georgia court so construes the statute, can it impose the same penalty under the same facts as could the Iowa courts under the reckless driving statute. On the other hand, if the Iowa court were to construe the Iowa statute as the Georgia court construes the Georgia statute, the punishment of the reckless driver in Iowa would be too severe.

In other words, the courts in these cases began with the results they desired to accomplish: they wanted to give what they considered to be adequate punishment to drunken drivers: their conclusions determined their reasoning.

But the conception that judges work back from conclusions to principles is so heretical that it seldom finds expression. Daily, judges, in connection with their decisions, deliver so-called opinions in which they purport to set forth the bases of their conclusions. Yet you will study these opinions in vain to discover anything remotely resembling a statement of the actual judging process. They are written in conformity with the time-honored theory. They picture the judge applying rules and principles to the facts, that is, taking some rule or principle

(usually derived from opinions in earlier cases) as his major premise, employing the facts of the case as the minor premise, and then coming to his judgment by processes of pure reasoning.

Now and again some judge, more clear-witted and outspoken than his fellows, describes (when off the bench) his methods in more homely terms. Recently Judge Hutcheson essayed such an honest report of the judicial process. He tells us that after canvassing all the available material at his command and duly cogitating on it, he gives his imagination play,

> and brooding over the case, waits for the feeling, the hunch—that intuitive flash of understanding that makes the jump-spark connection between question and decision and at the point where the path is darkest for the judicial feet, sets its light along the way. . . . In feeling or "hunching" out his decisions, the judge acts not differently from but precisely as the lawyers do in working on their cases, with only this exception, that the lawyer, in having a predetermined destination in view,—to win the law-suit for his client—looks for and regards only those hunches which keep him in the path that he has chosen, while the judge, being merely on his way with a roving commission to find the just solution, will follow his hunch wherever it leads him. . . .

And Judge Hutcheson adds:

> I must premise that I speak now of the judgment or decision, the solution itself, as opposed to the apologia for that decision; the decree, as opposed to the logomachy, the effusion of the judge by which that decree is explained or excused. . . . The judge really decides by feeling and not by judgment, by hunching and not by ratiocination, such ratiocination appearing only in the opinion. The vital motivating impulse for the decision is an intuitive sense of what is right or wrong in the particular case; and the astute judge, having so decided, enlists his every faculty and belabors his laggard mind, not only to justify that intuition to himself, but to make it pass muster with his critics. [Accordingly, he passes in review all of the rules, principles, legal categories, and concepts] which he may find useful, directly or by an analogy, so as to select from them those which in his opinion will justify his desired result.

We may accept this as an approximately correct description of how all judges do their thinking. But see the consequences. If the law consists of the decisions of the judges and if those decisions are based on the judge's hunches, then the way in which the judge gets his hunches is the key to the judicial process. Whatever produces the judge's hunches makes the law.

What, then, are the hunch-producers? What are the stimuli which make a judge feel that he should try to justify one conclusion rather than another?

The rules and principles of law are one class of such stimuli. But there are many others, concealed or unrevealed, not frequently considered in discussions of the character or nature of law. To the infrequent extent that these other stimuli have been considered at all, they have been usually referred to as "the political, economic and moral prejudices" of the judge. A moment's reflection would, indeed, induce any open-minded person to admit that factors of such character must be operating in the mind of the judge. . . .

ILLUSORY PRECEDENTS

Lawyers and judges purport to make large use of precedents; that is, they purport to rely on the conduct of judges in past cases as a means of procuring analogies for action in new cases. But since what was actually decided in the earlier cases is seldom revealed, it is impossible, in a real sense, to rely on these precedents. What the courts in fact do is to manipulate the language of former decisions. They could approximate a system of real precedents only if the judges, in rendering those former decisions, had reported with fidelity the precise steps by which they arrived at their decisions. The paradox of the situations is that, granting there is value in a system of precedents, our present use of illusory precedents makes the employment of real precedents impossible. . . .

The rules a judge announces when publishing his decision are, therefore, intelligible only if one can relive the judge's unique experience while he was trying the case—which, of course, cannot be done. . . .

Every lawyer of experience comes to know (more or less unconsciously) that in the great majority of cases, the precedents are none too good as bases of prediction. Somehow or other, there are

plenty of precedents to go around. A recent writer, a believer in the use of precedents, has said proudly that "it is very seldom indeed that a judge cannot find guidance of some kind, direct or indirect, in the mass of our reported decisions—by this time a huge accumulation of facts as well as rules." In plain English, . . . a court can usually find earlier decisions which can be made to appear to justify almost any conclusion.

DOING JUSTICE IN PARTICULAR CASES

What has just been said is not intended to mean that most courts arrive at their conclusions arbitrarily or apply a process of casuistical deception in writing their opinions. The process we have been describing involves no insincerity or duplicity. The average judge sincerely believes that he is using his intellect as "a cold logic engine" in applying rules and principles derived from the earlier cases to the objective facts of the case before him.

. . . Swayed by the belief that their opinions will serve as precedents and will therefore bind the thought processes of judges in cases which may thereafter arise, they feel obliged to consider excessively not only what has previously been said by other judges but also the future effect of those generalizations which they themselves set forth as explanations of their own decisions. When publishing the rules which are supposed to be the core of their decisions, they thus feel obligated to look too far both backwards and forwards. Many a judge, when unable to find old worn patterns which will fit his conclusions, is overcautious about announcing a so-called new rule for fear that, although the new rule may lead to a just conclusion in the case before him, it may lead to undesirable results in the future—that is, in cases not then before the court. Once trapped by the belief that the announced rules are the paramount thing in the law, and that uniformity and certainty are of major importance and are to be procured by uniformity and certainty in the phrasing of rules, a judge is likely to be affected, in determining what is fair to the parties in the unique situation before him, by consideration of the possible, yet scarcely imaginable, bad effect of a just opinion in the instant case on possible unlike cases which may later be brought into court. He then refuses to do justice in the case on trial because he fears that "hard cases make bad laws." And thus arises what may aptly be called "injustice according to law." . . .

The judge, at his best, is an arbitrator, a "sound man" who strives to do justice to the parties by exercising a wise discretion with reference to the peculiar circumstances of the case. He does not merely "find" or invent some generalized rule which he "applied" to the facts presented to him. He does "equity" in the sense in which Aristotle—when thinking most clearly—described it. "It is equity," he wrote in his Rhetoric, "to pardon human failings, and to look to the law giver and not to the law; . . . to prefer arbitration to judgment, for the arbitrator sees what is equitable, but the judge only the law, and for this an arbitrator was first appointed, in order that equity might flourish." The bench and bar usually try to conceal the arbitral function of the judge. . . . But although fear of legal uncertainty leads to this concealment, the arbitral function is the central fact in the administration of justice. . . .

WHAT THE LAW IS

. . . We may now venture a rough definition of law from the point of view of the average man: For any particular lay person, the law, with respect to any particular set of facts, is a decision of a court with respect to those facts so far as that decision affects that particular person. Until a court has passed on those facts no law on that subject is yet in existence. Prior to such a decision, the only law available is the opinion of lawyers as to the law relating to that person and to those facts. Such opinion is not actually law but only a guess as to what a court will decide.

Law, then, as to any given situation is either (a) actual law, *i.e.,* a specific past decision, as to that situation, or (b) probable law, *i.e.,* a guess as to a specific future decision.

Usually, when a client consults his lawyer about "the law," his purpose is to ascertain not what courts have actually decided in the past but what the courts will probably decide in the future. He asks, "Have I a right, as a stockholder of the American Taffy Company of Indiana, to look at

the corporate books?" Or, "Do I have to pay an inheritance tax to the State of New York on bonds left me by my deceased wife, if our residence was in Ohio, but the bonds, at the time of her death, were in a safety deposit box in New York?" Or, "Is there a right of 'peaceful' picketing in a strike in the State of California?" Or, "If Jones sells me his Chicago shoe business and agrees not to compete for ten years, will the agreement be binding?" The answers (although they may run "There is such a right," "The law is that the property is not taxable," "Such picketing is unlawful," "The agreement is not legally binding") are in fact prophecies or predictions of judicial action. It is from this point of view that the practice of law has been aptly termed an art of prediction. . . .

While the majority of lawyers deny that judges make law, a vigorous minority assert, realistically, that they do. But when does a judge make law? The minority here splits into two groups.

John Chipman Gray is typical of the first group. His contribution to hard-headed thinking about law was invaluable. He compelled his readers to differentiate between *law* and the *sources* of law. "The Law of the State," he wrote, "is composed of the rules which the courts, that is the judicial organs of that body, lay down for the determination of legal rights and duties." He felt it absurd to affirm the existence of law which the courts do not follow: "The Law of a State . . . is not an ideal, but something which actually exists." His thesis was that "the Law is made up of the rules for decision which the courts lay down; that all such rules are Law; that *rules for conduct which the courts do not apply are not Law; that the fact that courts apply rules is what makes them Law; that there is no mysterious entity 'The Law' apart from these rules; and that the judges are rather the creators than the discoverers of the Law.*"

According to Gray, the "law of a great nation" means "the opinions of a half-a-dozen old gentlemen. . . ." For, "if those half-a-dozen old gentlemen form the highest tribunal of a country, then no rule or principle which they refuse to follow is Law in that country." Of course, he added, "those six men seek the rules which they follow not in their own whims, but they derive them from *sources* . . . to which they are directed, by the organized body (the State) to which they belong, to apply themselves."

And those sources of law—*i.e.,* sources of "the rules for decision which the courts lay down"—are statutes, judicial precedents, opinions of experts, customs and principles of morality (using the term morality to include "public policy"). That none of these factors is, in and of itself, law is best exemplified by a consideration of a most important source—statutes. For, says Gray, after all it is only words that the legislature utters when it enacts a statute. And these words can get into action only through the rules laid down by the courts: it is for the courts to say what those words mean. There are limits to the courts' power of interpretation, but those limits are vague and undefined. And that is why statutes are not part of the Law itself, but only a source of law: "It has sometimes been said that the Law is composed of two parts—legislative law and judge-made law, but in truth all the Law is judge-made law. The shape in which a statute is imposed on the community as a guide for conduct is that statute as interpreted by the courts. The courts put life into the dead words of the statute. To quote . . . from Bishop Hoadly: 'Nay, whoever hath an *absolute authority* to *interpret* any written or spoken laws, it is *He* who is truly the Law Giver to all intents and purposes, and not the Person who first wrote and spoke them.'"

Gray was indeed a hardy foe of the . . . fundamentalists. Judges, he saw, make the law and, until they make it, there isn't any law, but only ingredients for making law. When a handful of old gentlemen who compose the highest court announce the law, that is the Law, until they change it, whether anyone else, however wise, thinks it good or bad, right or wrong. *But,* for all his terse directness, you will detect more than a trace of the old philosophy in Gray's views. You will note his constant reiteration of the words "rules" and "principles." Gray defines law not as what courts decide but as the "*rules* which the courts lay down for the determination of legal rights and duties" or "the *rules of decision* which the courts lay down." If a court in deciding a particular case fails to apply the "rule generally followed," that decision is not law. The rule for decisions usually laid down by the courts in Massachusetts is that a payment made on Sunday discharges a debt. "A judge in Massachusetts once decided that payment on Sunday was no discharge of a debt, but that has never been the Law of Massachusetts," said Gray.

Judges make law, according to Gray, when they make or change the rules; lawmaking is legal rule-making, the promulgation by a judge of a new rule for decision.

Now this stress on generality as the essence of law is a remnant of the old myth. And a vigorous remnant. It is found in the thinking of perhaps ninety percent of even those who, like Gray, scoff at the idea that law making occurs anywhere except in the court-room. Unless, they say, a court announces a new rule—announces it expressly or impliedly—it is not making law. Law equals legal rules—rules which the courts use, not anyone else's rules, but rules nevertheless; such judge-made rules constitute the law.

But in 1897 a new attitude was expressed when Holmes wrote, "A legal duty so called is nothing but a prediction that if a man does or omits certain things he will be made to suffer in this or that way by a judgment of the court; and so of a legal right. . . . *If you want to know the law and nothing else, you must look at it as a bad man, who cares only for the material consequences which such knowledge enables him to predict.* . . . What constitutes the law? You will find some text writers telling you that it is something different from what is decided by the courts of Massachusetts or England, that it is a system of reason, that it is a deduction from principles of ethics or admitted axioms or what not. But if we take the view of our friend the bad man we shall find that he does not care two straws for the axioms or deductions, but that he does want to know what the Massachusetts or English courts are likely to do in fact. I am much of his mind. *The prophecies of what the courts will do in fact, and nothing more pretentious, are what I mean by law.*"

That was in 1897. In 1899 Holmes said, "We must think things not words, or at least we must constantly translate our words into the facts for which they stand if we are to keep to the real and the true. I sometimes tell law students that the law schools pursue an inspirational method combined with a logical method, that is, the postulates are taken for granted upon authority without inquiry into their worth, and then logic is used as the only tool to develop the results. It is a necessary method for the purpose of teaching dogma. But inasmuch as the real justification of a rule of law, if there be one, is that it helps to bring about a social end which we desire, it is no less necessary that those who make and develop the law should have those ends articulately in their mind. . . . *A generalization is empty so far as it is general. Its value depends on the number of particulars which it calls up to the speaker and the bearer.*"

RULES AND LAW

Holmes's description of law can be stated as a revision of Gray's definition, thus: Law is made up not of rules for decisions laid down by the courts but of the decisions themselves. All such decisions are law. The fact that courts render these decisions makes them law. There is no mysterious entity apart from these decisions. If the judges in any case come to a "wrong" result and give forth a decision which is discordant with their own or any one else's rules, their decision is none the less law. The "law of a great nation" means the decisions of a handful of old gentlemen, and whatever they refuse to decide is not law. Of course those old gentlemen in deciding cases do not follow their own whims, but derive their views from many sources. And among those sources are not only statutes, precedents, customs and the like, but the rules which other courts have announced when deciding cases. Those rules are no more law than statutes are law. For, after all, rules are merely words and those words can get into action only through decisions; it is for the courts in deciding any case to say what the rules mean, whether those rules are embodied in a statute or in the opinion of some other court. The shape in which rules are imposed on the community is those rules as translated into concrete decisions. Your bad man doesn't care what the rules may be if the decisions are in his favor. He is not concerned with any mysterious entity such as the Law of Massachusetts which consists of the rules usually applied by the courts; he regards only what a very definite court decides in the very definite case in which he is involved. . . .

Often when a judge decides a case he simultaneously publishes an essay, called an opinion, explaining that he used an old rule or invented a new rule to justify his judgment. But no matter what he says, it is his decision which fixes the legal

positions of the litigants. If Judge Brilliant decides that Mr. Evasion must pay the federal government $50,000 for back taxes or that Mrs. Goneril is entitled to nothing under the will of her father, Mr. Lear, the contents of the judge's literary effusion makes not one iota of practical difference to Mr. Evasion or Mrs. Goneril. Opinion or no opinion, opinion-with-a-new-rule-announced or opinion-with-old-rules-proclaimed—it is all one to the parties whose contentions he adjudicated.

To be sure, this opinion may affect Judge Conformity who is later called on to decide the case of Rex vs. Humpty Dumpty. If Judge Brilliant in Mr. Evasion's case describes a new legal doctrine, his innovations may be *one* of the factors which actuates Judge Conformity to decide for Humpty Dumpty, if Judge Conformity thinks the facts in Humpty Dumpty's case are like those in Mr. Evasion's case. But—need it be reiterated?—the new doctrine will be but one of the factors actuating Judge Conformity.

The business of the judges is to decide particular cases. They, or some third person viewing their handiwork, may choose to generalize from these decisions, may claim to find common elements in the decisions in the cases of Fox vs. Grapes and Hee vs. Haw and describe the common elements as "rules." But those descriptions of alleged common elements are, at best, some aid to lawyers in guessing or bringing about future judicial conduct or some help to judges in settling other disputes. The rules will not directly decide any other cases in any given way, nor authoritatively compel the judges to decide those other cases in any given way; nor make it possible for lawyers to bring it about that the judges will decide any other cases in any given way, nor infallibly to predict how the judges will decide any other cases. Rules . . . are not law.

REVIEW AND DISCUSSION QUESTIONS

1. Describe the "conventional view" of law that Frank attacks. Do you agree or disagree that judges do not discover the law, they make it?
2. Why does he think precedents are "illusory"?
3. If legal rules and principles do not usually decide particular cases, how then *should* a responsible judge make a decision, according to the realist?
4. According to Frank, should a judge worry about the precedent that his decision sets for future cases?
5. Describe the distinction between "sources" of law and "law" according to Frank.
6. What, finally, is law according to realists?
7. Frank was an appeals court judge, not an ordinary criminal or civil judge hearing day-to-day cases. How might this fact have influenced his thinking?

CHAPTER 9

Morality and the Law

H. L. A. Hart is widely regarded as the most important modern defender of legal positivism. His 1961 book *The Concept of Law* did much to reinvigorate philosophical discussion of law and legal obligation, including the claim of natural law that some sort of essential connection exists between law on one hand and justice or morality on the other. But Hart was far from just a defender of Austinian positivism; indeed, he finds much to criticize in Austin's writings despite broadly sharing his positivistic conclusions. Ronald Dworkin was Hart's student at Oxford and eventually succeeded him as professor of jurisprudence there. Dworkin's criticisms of Hart and of positivism generally have also been widely discussed, along with his fascinating defense of a kind of natural law approach (shorn of Aquinas's or Blackstone's religious assumptions) and his insightful analysis of legal interpretation. In the final essay, Tony Honoré argues that morality depends on law because its requirements cannot be settled independently of law.

Positivism and the Separation of Law and Morals

H. L. A. Hart

Perhaps more than any single philosopher, H. L. A. Hart is responsible for the upsurge of interest in jurisprudence and philosophy of law during the past few decades. Hart's criticisms of Austin and his brilliant defense of his own positivistic theory have provoked much debate and interest in the subject. Beginning with an account of the historical importance of positivism, he weighs in this essay various criticisms of different aspects of the theory, including its insistence on the separation of law and morals, and Austin's claim that laws are the commands of the sovereign. He then examines the "distinctly American criticism" leveled by legal realists. Hart was professor of jurisprudence at Oxford University.

In this article I shall discuss and attempt to defend a view which Mr. Justice Holmes, among others, held and for which he and they have been much criticized. . . ¿Contemporary voices tell us we must recognize something obscured by the legal "positivists" whose day is now over: that there is a "point of intersection between law and morals," or that what *is* and what *ought* to be are somehow indissolubly fused or inseparable, though the positivists denied it. What do these phrases mean? Or rather which of the many things that they *could* mean, *do* they mean? Which of them do "positivists" deny and why is it wrong to do so?

1

I shall present the subject as part of the history of an idea. At the close of the eighteenth century and the beginning of the nineteenth the most earnest thinkers in England about legal and social problems and the architects of great reforms were the great utilitarians. Two of them, Bentham and Austin, constantly insisted on the need to distinguish, firmly and with the maximum of clarity, law as it is from law as it ought to be. This theme haunts their work, and they condemned the natural-law thinkers precisely because they had blurred this apparently simple but vital distinction. By contrast, at the present time in this country and to a lesser extent in England, this separation between law and morals is held to be superficial and wrong. Some critics have thought that it blinds men to the true nature of law and its roots in social life. Others have thought it not only intellectually misleading but corrupting in practice, at its worst apt to weaken resistance to state tyranny or absolutism, and at its best apt to bring law into disrespect. The nonpejorative name "legal positivism," like most terms which are used as missiles in intellectual battles, has come to stand for a baffling multitude of different sins. One of them is the sin, real or alleged, of insisting, as Austin and Bentham did, on the separation of law as it is and law as it ought to be.

How then has this reversal of the wheel come about? What are the theoretical errors in this distinction? Have the practical consequences of stressing the distinction as Bentham and Austin did been bad? . . .

Bentham and Austin were not dry analysts fiddling with verbal distinctions while cities burned, but were the vanguard of a movement which laboured with passionate intensity and much success to bring about a better society and better laws. Why then did they insist on the separation of law as it is and law as it ought to be? What did they mean? Let us first see what they said. Austin formulated the doctrine:

> The existence of law is one thing; its merit or demerit is another. Whether it be or be not is one enquiry; whether it be or be not conformable to an assumed standard, is a different enquiry. A law, which actually exists, is a law, though we happen to dislike it, or though it vary from the text, by which we regulate our approbation and disapprobation. This truth, when formally announced as an abstract proposition, is so simple and glaring that it seems idle to insist upon it. But simple and glaring as it is, when enunciated in abstract expressions the enumeration of the instances in which it has been forgotten would fill a volume.
>
> Sir William Blackstone, for example, says in his "Commentaries," that the laws of God are superior in obligation to all other laws; that no human laws should be suffered to contradict them; that human laws are of no validity if contrary to them; and that all valid laws derive their force from that Divine original.
>
> Now, he *may* mean that all human laws ought to conform to the Divine laws. If this be his meaning, I assent to it without hesitation. . . . Perhaps, again, he means that human lawgivers are themselves obliged by the Divine laws to fashion the laws which they impose by that ultimate standard, because if they do not, God will punish them. To this also I entirely assent. . . .
>
> But the meaning of this passage of Blackstone, if it has a meaning, seems rather to be this: that no human law which conflicts with the Divine law is obligatory or binding; in other words, that no human law which conflicts with the Divine law *is a law*. . . .

Austin's protest against blurring the distinction between what law is and what it ought to be is quite general: it is a mistake, whatever our standard of what ought to be, whatever "the text by which we regulate our approbation or disapprobation." His examples, however, are always a confusion between law as it is and law as morality would require it to be. For him, it must be remembered,

the fundamental principles of morality were God's commands, to which utility was an "index": besides this there was the actual accepted morality of a social group or "positive" morality. . . .

In view of later criticisms it is also important to distinguish several things that the [positivist] utilitarians did not mean by insisting on their separation of law and morals. They certainly accepted many of the things that might be called "the intersection of law and morals." First, they never denied that, as a matter of historical fact, the development of legal systems had been powerfully influenced by moral opinion, and, conversely, that moral standards had been profoundly influenced by law, so that the content of many legal rules mirrored moral rules or principles. . . .

Secondly, neither Bentham nor his followers denied that by explicit legal provisions moral principles might at different points be brought into a legal system and form part of its rules, or that courts might be legally bound to decide in accordance with what they thought just or best. . . .

What both Bentham and Austin were anxious to assert were the following two simple things: first, in the absence of an expressed constitutional or legal provision, it could not follow from the mere fact that a rule violated standards of morality that it was not a rule of law; and, conversely, it could not follow from the mere fact that a rule was morally desirable that it was a rule of law. . . .

2

So much for the [positivist] doctrine in the heyday of its success. Let us turn now to some of the criticisms.

. . . We must remember that the [positivist] utilitarians combined with their insistence on the separation of law and morals two other equally famous but distinct doctrines. One was the important truth that a purely analytical study of legal concepts, a study of the meaning of the distinctive vocabulary of the law, was as vital to our understanding of the nature of law as historical or sociological studies, though of course it could not supplant them. The other doctrine was the famous imperative theory of law—that law is essentially a command.

These three doctrines constitute the [positivist] tradition in jurisprudence; yet they are distinct doctrines. It is possible to endorse the separation between law and morals and to value analytical inquiries into the meaning of legal concepts and yet think it wrong to conceive of law as essentially a command. . . . [Some] critics . . . have thought that the inadequacies of the command theory which gradually came to light were sufficient to demonstrate the falsity of the separation of law and morals.

This was a mistake, but a natural one. To see how natural it was we must look a little more closely at the command idea. . . . The [positivist] utilitarians thought that the essence of a legal system could be conveyed if the notion of a command were supplemented by that of a habit of obedience. The simple scheme was this: What is a command? It is simply an expression by one person of the desire that another person should do or abstain from some action, accompanied by a threat of punishment which is likely to follow disobedience. Commands are laws if two conditions are satisfied: First, they must be general; second, they must be commanded by what (as both Bentham and Austin claimed) exists in every political society whatever its constitutional form, namely, a person or a group of persons who are in receipt of habitual obedience from most of the society but pay no such obedience to others. These persons are its sovereign. Thus law is the command of the uncommanded commanders of society—the creation of the legally untrammelled will of the sovereign who is by definition outside the law.

It is easy to see that this account of a legal system is threadbare. One can also see why it might seem that its inadequacy is due to the omission of some essential connection with morality. The situation which the simple trilogy of command, sanction, and sovereign avails to describe, if you take these notions at all precisely, is like that of a gunman saying to his victim, "Give me your money or your life." The only difference is that in the case of a legal system the gunman says it to a large number of people who are accustomed to the racket and habitually surrender to it. Law surely is not the gunman situation writ large, and legal order is surely not to be thus simply identified with compulsion.

This scheme, despite the points of obvious analogy between a statute and a command, omits some of the most characteristic elements of law. Let me

cite a few. It is wrong to think of a legislature (and a fortiori an electorate) with a changing membership, as a group of persons habitually obeyed: this simple idea is suited only to a monarch sufficiently long-lived for a "habit" to grow up. Even if we waive this point, nothing which legislators do makes law unless they comply with fundamental accepted rules specifying the essential lawmaking procedures. This is true even in a system having a simple unitary constitution like the British. These fundamental accepted rules specifying what the legislature must do to legislate are not commands habitually obeyed, nor can they be expressed as habits of obedience to persons. They lie at the root of a legal system, and what is most missing in the utilitarian scheme is an analysis of what it is for a social group and its officials to accept such rules. This notion, not that of a command as Austin claimed, is the "key to the science of jurisprudence," or at least one of the keys.

Again, Austin, in the case of a democracy, looked past the legislators to the electorate as "the sovereign" (or in England as part of it). He thought that in the United States the mass of the electors to the state and federal legislatures were the sovereign whose commands, given by their "agents" in the legislatures, were law. But on this footing the whole notion of the sovereign outside the law being "habitually obeyed" by the "bulk" of the population must go: for in this case the "bulk" obeys the bulk, that is, it obeys itself. Plainly the general acceptance of the authority of a lawmaking procedure, irrespective of the changing individuals who operate it from time to time, can be only distorted by an analysis in terms of mass habitual obedience to certain persons who are by definition outside the law, just as the cognate but much simpler phenomenon of the general social acceptance of a rule, say of taking off the hat when entering a church, would be distorted if represented as habitual obedience by the mass to specific persons.

Other critics dimly sensed a further and more important defect in the command theory, yet blurred the edge of an important criticism by assuming that the defect was due to the failure to insist upon some important connection between law and morals. This more radical defect is as follows. The picture that the command theory draws of life under law is essentially a simple relationship of the commander to the commanded, of superior to inferior, of top to bottom; the relationship is vertical between the commanders or authors of the law conceived of as essentially outside the law and those who are commanded and subject to the law. In this picture no place, or only an accidental or subordinate place, is afforded for a distinction between types of legal rules which are in fact radically different. Some laws require men to act in certain ways or to abstain from acting whether they wish to or not. The criminal law consists largely of rules of this sort: like commands they are simply "obeyed" or "disobeyed." But other legal rules are presented to society in quite different ways and have quite different functions. They provide facilities more or less elaborate for individuals to create structures of rights and duties for the conduct of life within the coercive framework of the law. Such are the rules enabling individuals to make contracts, wills, and trusts, and generally to mould their legal relations with others. Such rules, unlike the criminal law, are not factors designed to obstruct wishes and choices of an antisocial sort. On the contrary, these rules provide facilities for the realization of wishes and choices. They do not say (like commands) "do this whether you wish it or not," but rather "if you wish to do this, here is the way to do it." Under these rules we exercise powers, make claims, and assert rights. These phrases mark off characteristic features of laws that confer rights and powers; they are laws which are, so to speak, put at the disposition of individuals in a way in which the criminal law is not.

. . . Rules that confer rights, though distinct from commands, need not be moral rules or coincide with them. Rights, after all, exist under the rules of ceremonies, games, and in many other spheres regulated by rules which are irrelevant to the question of justice or what the law ought to be. Nor need rules which confer rights be just or morally good rules. The rights of a master over his slaves show us that. "Their merit or demerit," as Austin termed it, depends on how rights are distributed in society and over whom or what they are exercised. These critics indeed revealed the inadequacy of the simple notions of command and habit for the analysis of law; at many points it is apparent that the social acceptance of a rule or standard of authority (even if it is motivated only by fear or

superstition or rests on inertia) must be brought into the analysis and cannot itself be reduced to the two simple terms. Yet nothing in this showed the utilitarian insistence on the distinction between the existence of law and its "merits" to be wrong.

3

I now turn to a distinctively American criticism of the separation of the law that is from the law that ought to be. It emerged from the critical study of the judicial process with which American jurisprudence has been on the whole so beneficially occupied. The most skeptical of these critics—the loosely named "Realists" of the 1930s—perhaps too naïvely accepted the conceptual framework of the natural sciences as adequate for the characterization of law and for the analysis of rule-guided action of which a living system of law at least partly consists. But they opened men's eyes to what actually goes on when courts decide cases, and the contrast they drew between the actual facts of judicial decision and the traditional terminology for describing it as if it were a wholly logical operation was usually illuminating; for in spite of some exaggeration the "Realists" made us acutely conscious of one cardinal feature of human language and human thought, emphasis on which is vital not only for the understanding of law but in areas of philosophy far beyond the confines of jurisprudence. The insight of this school may be presented in the following example. A legal rule forbids you to take a vehicle into the public park. Plainly this forbids an automobile, but what about bicycles, roller skates, toy automobiles? What about airplanes? Are these, as we say, to be called "vehicles" for the purpose of the rule or not? If we are to communicate with each other at all, and if, as in the most elementary form of law, we are to express our intentions that a certain type of behavior be regulated by rules, then the general words we use—like "vehicle" in the case I consider—must have some standard instance in which no doubts are felt about its application. There must be a core of settled meaning, but there will be, as well, a penumbra of debatable cases in which words are neither obviously applicable nor obviously ruled out. These cases will each have some features in common with the standard case; they will lack others or be accompanied by features not present in the standard case. Human invention and natural processes continually throw up such variants on the familiar, and if we are to say that these ranges of facts do or do not fall under existing rules, then the classifier must make a decision which is not dictated to him, for the facts and phenomena to which we fit our words and apply our rules are as it were *dumb.* The toy automobile cannot speak up and say, "I am a vehicle for the purpose of this legal rule," nor can the roller skates chorus, "We are not a vehicle." Fact situations do not await us neatly labeled, creased, and folded, nor is their legal classification written on them to be simply read off by the judge. Instead, in applying legal rules, someone must take the responsibility of deciding that words do or do not cover some case in hand with all the practical consequences involved in this decision.

We may call the problems which arise outside the hard core of standard instances or settled meaning "problems of the penumbra"; they are always with us whether in relation to such trivial things as the regulation of the use of the public park or in relation to the multidimensional generalities of a constitution. If a penumbra of uncertainty must surround all legal rules, then their application to specific cases in the penumbral area cannot be a matter of logical deduction, and so deductive reasoning, which for generations has been cherished as the very perfection of human reasoning, cannot serve as a model for what judges, or indeed anyone, should do in bringing particular cases under general rules. In this area men cannot live by deduction alone. And it follows that if legal arguments and legal decisions of penumbral questions are to be rational, their rationality must lie in something other than a logical relation to premises. So if it is rational or "sound" to argue and to decide that for the purposes of this rule an airplane is not a vehicle, this argument must be sound or rational without being logically conclusive. What is it then that makes such decisions correct or at least better than alternative decisions? Again, it seems true to say that the criterion which makes a decision sound in such cases is some concept of what the law ought to be; it is easy to slide from that into saying that it must be a moral judgment about what law ought to be. So here we touch upon a point of necessary

"intersection between law and morals" which demonstrates the falsity or, at any rate, the misleading character of the utilitarians' emphatic insistence on the separation of law as it is and ought to be. Surely, Bentham and Austin could only have written as they did because they misunderstood or neglected this aspect of the judicial process, because they ignored the problems of the penumbra.

The misconception of the judicial process which ignores the problems of the penumbra and which views the process as consisting preeminently in deductive reasoning is often stigmatized as the error of "formalism" or "literalism." My question now is, how and to what extent does the demonstration of this error show the [positivist] distinction to be wrong or misleading? Here there are many issues which have been confused, but I can only disentangle some. The charge of formalism has been leveled both at the "positivist" legal theorist and at the courts, but of course it must be a very different charge in each case. Leveled at the legal theorist, the charge means that he has made a theoretical mistake about the character of legal decision; he has thought of the reasoning involved as consisting in deduction from premises in which the judges' practical choices or decision play no part. It would be easy to show that Austin was guiltless of this error; only an entire misconception of what analytical jurisprudence is and why he thought it important has led to the view that he, or any other analyst, believed that the law was a closed logical system in which judges deduced their decisions from premises. On the contrary, he was very much alive to the character of language, to its vagueness or open character; he thought that in the penumbral situation judges must necessarily legislate, and, in accents that sometimes recall those of the late Judge Jerome Frank, he berated the common-law judges for legislating feebly and timidly and for blindly relying on real or fancied analogies with past cases instead of adapting their decisions to the growing needs of society as revealed by the moral standard of utility. The villains of this piece, responsible for the conception of the judge as an automaton, are not the utilitarian thinkers. The responsibility, if it is to be laid at the door of any theorist, is with . . . Blackstone's "childish fiction" (as Austin termed it) that judges only "find," never "make," law.

But we are concerned with "formalism" as a vice not of jurists but of judges. What precisely is it for a judge to commit this error, to be a "formalist," "automatic," a "slot machine"? . . . It is clear that the essence of his error is to give some general term an interpretation which is blind to social values and consequences (or which is in some other way stupid or perhaps merely disliked by critics). But logic does not prescribe interpretation of terms; it dictates neither the stupid nor intelligent interpretation of any expression. Logic only tells you hypothetically that *if* you give a certain term a certain interpretation then a certain conclusion follows. Logic is silent on how to classify particulars—and this is the heart of a judicial decision. So this reference to logic and to logical extremes is a misnomer for something else, which must be this. A judge has to apply a rule to a concrete case—perhaps the rule that one may not take a stolen "vehicle" across state lines, and in this case an airplane has been taken. He either does not see or pretends not to see that the general terms of this rule are susceptible of different interpretations and that he has a choice left open uncontrolled by linguistic conventions. He ignores, or is blind to, the fact that he is in the area of the penumbra and is not dealing with a standard case. Instead of choosing in the light of social aims, the judge fixes the meaning in a different way. He either takes the meaning that the word most obviously suggests in its ordinary nonlegal context to ordinary men, or one which the word has been given in some other legal context, or, still worse, he thinks of a standard case and then arbitrarily identifies certain features in it—for example, in the case of a vehicle, (1) normally used on land, (2) capable of carrying a human person, (3) capable of being self-propelled—and treats these three as always necessary and always sufficient conditions for the use in all contexts of the word "vehicle," irrespective of the social consequences of giving it this interpretation. This choice, not "logic," would force the judge to include a toy motor car (if electrically propelled) and to exclude bicycles and the airplane. In all this there is possibly great stupidity but not more "logic," and no less, than in cases in which the interpretation given to a general term and the consequent application of some general rule to a particular case is consciously controlled by some identified social aim.

Decisions made in a fashion as blind as this would scarcely deserve the name of decisions; we might as well toss a penny in applying a rule of law. But it is at least doubtful whether any judicial decisions (even in England) have been quite as automatic as this. Rather either the interpretations stigmatized as automatic have resulted from the conviction that it is fairer in a criminal statute to take a meaning which would jump to the mind of the ordinary man at the cost even of defeating other values, and this itself is a social policy (though possibly a bad one); or much more frequently, what is stigmatized as "mechanical" and "automatic" is a determined choice made indeed in the light of a social aim but of a conservative social aim. Certainly many of the Supreme Court decisions at the turn of the century which have been so stigmatized represent clear choices in the penumbral area to give effect to a policy of a conservative type. . . .

But how does the wrongness of deciding cases in an automatic and mechanical way and the rightness of deciding cases by reference to social purposes show that the [positivist] insistence on the distinction between what the law is and what it ought to be is wrong? I take it that no one who wished to use these vices of formalism as proof that the distinction between what is and what ought to be is mistaken would deny that the decisions stigmatized as automatic are law; nor would he deny that the system in which such automatic decisions are made is a legal system. Surely he would say that they are law, but they are bad law, they ought not to be law. But this would be to use the distinction, not to refute it; and of course both Bentham and Austin used it to attack judges for failing to decide penumbral cases in accordance with the growing needs of society.

Clearly, if the demonstration of the errors of formalism is to show the [positivist] distinction to be wrong, the point must be drastically restated. The point must be not merely that a judicial decision to be rational must be made in the light of some conception of what ought to be, but that the aims, the social policies and purposes to which judges should appeal if their decisions are to be rational, are themselves to be considered as part of the law in some suitably wide sense of "law" which is held to be more illuminating than that used by the utilitarians. This restatement of the point would have the following consequence: Instead of saying that the recurrence of penumbral questions shows us that legal rules are essentially incomplete, and that, when they fail to determine decisions, judges must legislate and so exercise a creative choice between alternatives, we shall say that the social policies which guide the judges' choice are in a sense there for them to discover; the judges are only "drawing out" of the rule what, if it is properly understood, is "latent" within it. To call this judicial legislation is to obscure some essential continuity between the clear cases of the rule's application and the penumbral decisions. I shall question later whether this way of talking is salutary, but I wish at this time to point out something obvious, but likely, if not stated, to tangle the issues. It does not follow that, because the opposite of a decision reached blindly in the formalist or literalist manner is a decision intelligently reached by reference to some conception of what ought to be, we have a junction of law and morals. . . . The point here is that intelligent decisions which we oppose to mechanical or formal decisions are not necessarily identical with decisions defensible on moral grounds. We may say of many a decision: "Yes, that is right; that is as it ought to be," and we may mean only that some accepted purpose or policy has been thereby advanced; we may not mean to endorse the moral propriety of the policy or the decision. So the contrast between the mechanical decision and the intelligent one can be reproduced inside a system dedicated to the pursuit of the most evil aims. . . .

We can now return to the main point. It is true that the intelligent decision of penumbral questions is one made not mechanically but in the light of aims, purposes, and policies, though not necessarily in the light of anything we would call moral principles. But is it wise to express this important fact by saying that the firm utilitarian distinction between what the law is and what it ought to be should be dropped? Perhaps the claim that it is wise cannot be theoretically refuted for it is, in effect, an *invitation* to revise our conception of what a legal rule is. We are invited to include in the "rule" the various aims and policies in the light of which its penumbral cases are decided on the ground that these aims have, because of their importance, as much right to be called law as the core of legal rules whose meaning is settled. But though an invitation cannot be refuted, it may be refused and I would proffer two

reasons for refusing this invitation. First, every thing we have learned about the judicial process can be expressed in other less mysterious ways. We can say laws are incurably incomplete and we must decide the penumbral cases rationally by reference to social aims. I think Holmes, who had such a vivid appreciation of the fact that "general propositions do not decide concrete cases," would have put it that way. Second, to insist on the utilitarian distinction is to emphasize that the hard core of settled meaning is law in some centrally important sense and that even if there are borderlines, there must first be lines. If this were not so the notion of rules controlling courts' decisions would be senseless as some of the "Realists"—in their most extreme moods, and, I think, on bad grounds—claimed.

By contrast, to soften the distinction, to assert mysteriously that there is some fused identity between law as it is and as it ought to be, is to suggest that all legal questions are fundamentally like those of the penumbra. It is to assert that there is no central element of actual law to be seen in the core of central meaning which rules have, that there is nothing in the nature of a legal rule inconsistent with *all* questions being open to reconsideration in the light of social policy. Of course, it is good to be occupied with the penumbra. Its problems are rightly the daily diet of the law schools. But to be occupied with the penumbra is one thing, to be preoccupied with it another. And preoccupation with the penumbra is, if I may say so, as rich a source of confusion in the American legal tradition as formalism in the English. Of course we might abandon the notion that rules have authority; we might cease to attach force or even meaning to an argument that a case falls clearly within a rule and the scope of a precedent. We might call all such reasoning "automatic" or "mechanical," which is already the routine invective of the courts. But until we decide that this *is* what we want, we should not encourage it by obliterating the utilitarian distinction. . . .

4

The third criticism of the separation of law and morals is of a very different character: it certainly is less an intellectual argument against the utilitarian distinction than a passionate appeal supported not by detailed reasoning but by reminders of a terrible experience. For it consists of the testimony of those who have descended into Hell, and, like Ulysses or Dante, brought back a message for human beings. Only in this case the Hell was not beneath or beyond earth, but on it; it was a Hell created on earth by men for other men.

This appeal comes from those German thinkers who lived through the Nazi regime and reflected upon its evil manifestations in the legal system. One of these thinkers, Gustav Radbruch, had himself shared the "positivist" doctrine until the Nazi tyranny, but he was converted by this experience and so his appeal to other men to discard the doctrine of the separation of law and morals has the special poignancy of a recantation. What is important about this criticism is that it really does confront the particular point which Bentham and Austin had in mind in urging the separation of law as it is and as it ought to be. . . .

. . . Austin, it may be recalled, was emphatic in condemning those who said that if human laws conflicted with the fundamental principles of morality that they cease to be laws, as talking "stark nonsense."

> The most pernicious laws, and therefore those which are most opposed to the will of God, have been and are continually enforced as laws by judicial tribunals. Suppose an act innocuous, or positively beneficial, be prohibited by the sovereign under the penalty of death; if I commit this act, I shall be tried and condemned, and if I object to the sentence, that it is contrary to the law of God . . . the court of justice will demonstrate the inconclusiveness of my reasoning by hanging me up, in pursuance of the law of which I have impugned the validity. An exception, demurrer, or plea, founded on the law of God was never heard in a Court of Justice, from the creation of the world down to the present moment.

These are strong, indeed brutal words, but we must remember that they went along—in the case of Austin and, of course, Bentham—with the conviction that if laws reached a certain degree of iniquity then there would be a plain moral obligation to resist them and to withhold obedience. We shall see, when we consider the alternatives, that this simple presentation of the human dilemma which may arise has much to be said for it. . . .

It is impossible to read without sympathy [the] passionate demand that the German legal conscience should be open to the demands of morality and [the] complaint that this has been too little the case in the German tradition. On the other hand there is an extraordinary naïveté in the view that insensitiveness to the demands of morality and subservience to state power in a people like the Germans should have arisen from the belief that law might be law though it failed to conform with the minimum requirements of morality. Rather this terrible history prompts inquiry into why emphasis on the slogan "law is law," and the distinction between law and morals, acquired a sinister character in Germany, but elsewhere, as with the utilitarians themselves, went along with the most enlightened liberal attitudes.

. . . If with the [positivist] utilitarians we speak plainly, we say that laws may be law but too evil to be obeyed. This is a moral condemnation which everyone can understand and it makes an immediate and obvious claim to moral attention. If, on the other hand, we formulate our objection as an assertion that these evil things are not law, here is an assertion which many people do not believe, and if they are disposed to consider it at all, it would seem to raise a whole host of philosophical issues before it can be accepted. So perhaps the most important single lesson to be learned from this form of the denial of the [positivist] distinction is the one that the utilitarians were most concerned to teach: when we have the ample resources of plain speech we must not present the moral criticism of institutions as propositions of a disputable philosophy.

REVIEW AND DISCUSSION QUESTIONS

1. What are the key claims made by legal positivists ("utilitarians" as Hart sometimes calls them)?
2. Why does Hart reject Austin's "command" theory of law?
3. What is the distinction between the "core" and the "penumbra," and how is it relevant to the dispute between formalism and legal realism?
4. How does Hart answer those who suggest that legal positivism contributed to the Germans' willingness to accept Nazism?
5. How might Austin respond to Hart's criticisms?

Law as the Union of Primary and Secondary Rules

H. L. A. Hart

In the following selection taken from his influential book *The Concept of Law,* Hart presents his view of the nature of law, a theory that although still consistent with the basic tenets of positivism, nonetheless is not open to the objections raised against Austin. Hart begins by asking the reader to imagine the weaknesses of a society with only "primary" rules of obligation—it would suffer from uncertainty, it would be inefficient, and its rules would be static—and then goes on to explain why in a developed system it is essential that "secondary" rules be added to remedy these weaknesses. Law, for Hart, is an affair of rules, and the most important of these is the rule of recognition, which is the ultimate source of legal validity. He concludes by contrasting internal and external statements about the rule of recognition and distinguishing statements of validity from statements of value.

It is, of course, possible to imagine a society without a legislature, courts or officials of any kind. Indeed, there are many studies of primitive communities which not only claim that this possibility is realized but depict in detail the life of a society where the only means of social control is that general attitude of the group towards its own standard modes of behaviour in terms of which we have

characterized rules of obligation. A social structure of this kind is often referred to as one of "custom"; but we shall not use this term, because it often implies that the customary rules are very old and supported with less social pressure than other rules. To avoid these implications we shall refer to such a social structure as one of primary rules of obligation. If a society is to live by such primary rules alone, there are certain conditions which, granted a few of the most obvious truisms about human nature and the world we live in, must clearly be satisfied. The first of these conditions is that the rules must contain in some form restrictions on the free use of violence, theft, and deception to which human beings are tempted but which they must, in general, repress, if they are to co-exist in close proximity to each other. Such rules are in fact always found in the primitive societies of which we have knowledge, together with a variety of others imposing on individuals various positive duties to perform services or make contributions to the common life. Secondly, though such a society may exhibit the tension, already described, between those who accept the rules and those who reject the rules except where fear of social pressure induces them to conform, it is plain that the latter cannot be more than a minority, if so loosely organized a society of persons, approximately equal in physical strength, is to endure: for otherwise those who reject the rules would have too little social pressure to fear. This too is confirmed by what we know of primitive communities where, though there are dissidents and malefactors, the majority live by the rules seen from the internal point of view.

More important for our present purpose is the following consideration. It is plain that only a small community closely knit by ties of kinship, common sentiment, and belief, and placed in a stable environment, could live successfully by such a régime of unofficial rules. In any other conditions such a simple form of social control must prove defective and will require supplementation in different ways. In the first place, the rules by which the group lives will not form a system, but will simply be a set of separate standards, without any identifying or common mark, except of course that they are the rules which a particular group of human beings accepts. They will in this respect resemble our own rules of etiquette. Hence if doubts arise as to what the rules are or as to the precise scope of some given rule, there will be no procedure for settling this doubt, either by reference to an authoritative text or to an official whose declarations on this point are authoritative. For, plainly, such a procedure and the acknowledgment of either authoritative text or persons involve the existence of rules of a type different from the rules of obligation or duty which *ex hypothesi* are all that the group has. This defect in the simple social structure of primary rules we may call its *uncertainty.*

A second defect is the *static* character of the rules. The only mode of change in the rules known to such a society will be the slow process of growth, whereby courses of conduct once thought optional become first habitual or usual, and then obligatory, and the converse process of decay, when deviations, once severely dealt with, are first tolerated and then pass unnoticed. There will be no means, in such a society, of deliberately adapting the rules to changing circumstances, either by eliminating old rules or introducing new ones: for, again, the possibility of doing this presupposes the existence of rules of a different type from the primary rules of obligation by which alone the society lives. In an extreme case the rules may be static in a more drastic sense. This, though never perhaps fully realized in any actual community, is worth considering because the remedy for it is something very characteristic of law. In this extreme case, not only would there be no way of deliberately changing the general rules, but the obligations which arise under the rules in particular cases could not be varied or modified by the deliberate choice of any individual. Each individual would simply have fixed obligations or duties to do or abstain from doing certain things. It might indeed very often be the case that others would benefit from the performance of these obligations; yet if there are only primary rules of obligation they would have no power to release those bound from performance or to transfer to others the benefits which would accrue from performance. For such operations of release or transfer create changes in the initial positions of individuals under the primary rules of obligation, and for these operations to be possible there must be rules of a sort different from the primary rules.

The third defect of this simple form of social life is the *inefficiency* of the diffuse social pressure by

which the rules are maintained. Disputes as to whether an admitted rule has or has not been violated will always occur and will, in any but the smallest societies, continue interminably, if there is no agency specially empowered to ascertain finally, and authoritatively, the fact of violation. Lack of such final and authoritative determinations is to be distinguished from another weakness associated with it. This is the fact that punishments for violations of the rules, and other forms of social pressure involving physical effort or the use of force, are not administered by a special agency but are left to the individuals affected or to the group at large. It is obvious that the waste of time involved in the group's unorganized efforts to catch and punish offenders, and the smouldering vendettas which may result from self help in the absence of an official monopoly of "sanctions," may be serious. The history of law does, however, strongly suggest that the lack of official agencies to determine authoritatively the fact of violation of the rules is a much more serious defect; for many societies have remedies for this defect long before the other.

The remedy for each of these three main defects in this simplest form of social structure consists in supplementing the *primary* rules of obligation with *secondary* rules which are rules of a different kind. The introduction of the remedy for each defect might, in itself, be considered a step from the pre-legal into the legal world; since each remedy brings with it many elements that permeate law: certainly all three remedies together are enough to convert the régime of primary rules into what is indisputably a legal system. We shall consider in turn each of these remedies and show why law may most illuminatingly be characterized as a union of primary rules of obligation with such secondary rules. Before we do this, however, the following general points should be noted. Though the remedies consist in the introduction of rules which are certainly different from each other, as well as from the primary rules of obligation which they supplement, they have important features in common and are connected in various ways. Thus they may all be said to be on a different level from the primary rules, for they are all *about* such rules; in the sense that while primary rules are concerned with the actions that individuals must or must not do, these secondary rules are all concerned with the primary rules themselves. They specify the ways in which the primary rules may be conclusively ascertained, introduced, eliminated, varied, and the fact of their violation conclusively determined.

The simplest form of remedy for the *uncertainty* of the régime of primary rules is the introduction of what we shall call a "rule of recognition." This will specify some feature or features possession of which by a suggested rule is taken as a conclusive affirmative indication that it is a rule of the group to be supported by the social pressure it exerts. The existence of such a rule of recognition may take any of a huge variety of forms, simple or complex. It may, as in the early law of many societies, be no more than that an authoritative list or text of the rules is to be found in a written document or carved on some public monument. No doubt as a matter of history this step from the pre-legal to the legal may be accomplished in distinguishable stages, of which the first is the mere reduction to writing of hitherto unwritten rules. This is not itself the crucial step, though it is a very important one: what is crucial is the acknowledgement of reference to the writing or inscription as *authoritative,* i.e., as the *proper* way of disposing of doubts as to the existence of the rule. Where there is such an acknowledgement there is a very simple form of secondary rule: a rule for conclusive identification of the primary rules of obligation.

In a developed legal system the rules of recognition are of course more complex; instead of identifying rules exclusively by reference to a text or list they do so by reference to some general characteristic possessed by the primary rules. This may be the fact of their having been enacted by a specific body, or their long customary practice, or their relation to judicial decisions. Moreover, where more than one of such general characteristics are treated as identifying criteria, provision may be made for their possible conflict by their arrangement in an order of superiority, as by the common subordination of custom or precedent to statute, the latter being a "superior source" of law. Such complexity may make the rules of recognition in a modern legal system seem very different from the simple acceptance of an authoritative text: yet even in this simplest form, such a rule brings with it many elements distinctive of law. By providing an authoritative mark it introduces, although in

embryonic form, the idea of a legal system: for the rules are now not just a discrete unconnected set but are, in a simple way, unified. Further, in the simple operation of identifying a given rule as possessing the required feature of being an item on an authoritative list of rules we have the germ of the idea of legal validity.

The remedy for the *static* quality of the régime of primary rules consists in the introduction of what we shall call "rules of change." The simplest form of such a rule is that which empowers an individual or body of persons to introduce new primary rules for the conduct of the life of the group, or of some class within it, and to eliminate old rules. As we have already argued . . . it is in terms of such a rule, and not in terms of orders backed by threats, that the ideas of legislative enactment and repeal are to be understood. Such rules of change may be very simple or very complex: the powers conferred may be unrestricted or limited in various ways: and the rules may, besides specifying the persons who are to legislate, define in more or less rigid terms the procedure to be followed in legislation. Plainly, there will be a very close connexion between the rules of change and the rules of recognition: for where the former exists the latter will necessarily incorporate a reference to legislation as an identifying feature of the rules, though it need not refer to all the details of procedure involved in legislation. Usually some official certificate or official copy will, under the rules of recognition, be taken as a sufficient proof of due enactment. Of course if there is a social structure so simple that the only "source of law" is legislation, the rule of recognition will simply specify enactment as the unique identifying mark or criterion of validity of the rules. This will be the case for example in the imaginary kingdom of Rex I . . . [where] the rule of recognition would simply be that whatever Rex I enacts is law.

We have already described in some detail the rules which confer on individuals power to vary their initial positions under the primary rules. Without such private power-conferring rules society would lack some of the chief amenities which law confers upon it. For the operations which these rules make possible are the making of wills, contracts, transfers of property, and many other voluntarily created structures of rights and duties which typify life under law, though of course an elementary form of power-conferring rule also underlies the moral institution of a promise. The kinship of these rules with the rules of change involved in the notion of legislation is clear, and as recent theory such as Kelsen's has shown, many of the features which puzzle us in the institutions of contract or property are clarified by thinking of the operations of making a contract or transferring property as the exercise of limited legislative powers by individuals.

The third supplement to the simple régime of primary rules, intended to remedy the *inefficiency* of its diffused social pressure, consists of secondary rules empowering individuals to make authoritative determinations of the question whether, on a particular occasion, a primary rule has been broken. The minimal form of adjudication consists in such determinations, and we shall call the secondary rules which confer the power to make them "rules of adjudication." Besides identifying the individuals who are to adjudicate, such rules will also define the procedure to be followed. Like the other secondary rules these are on a different level from the primary rules: though they may be reinforced by further rules imposing duties on judges to adjudicate, they do not impose duties but confer judicial powers and a special status on judicial declarations about the breach of obligations. Again these rules, like the other secondary rules, define a group of important legal concepts: in this case the concepts of judge or court, jurisdiction and judgment. Besides these resemblances to the other secondary rules, rules of adjudication have intimate connexions with them. Indeed, a system which has rules of adjudication is necessarily also committed to a rule of recognition of an elementary and imperfect sort. This is so because, if courts are empowered to make authoritative determinations of the fact that a rule has been broken, these cannot avoid being taken as authoritative determinations of what the rules are. So the rule which confers jurisdiction will also be a rule of recognition, identifying the primary rules through the judgments of the courts and these judgments will become a "source" of law. It is true that this form of rule of recognition, inseparable from the minimum form of jurisdiction, will be very imperfect. Unlike an authoritative text or a statute book, judgments may not be couched in general terms

and their use as authoritative guides to the rules depends on a somewhat shaky inference from particular decisions, and the reliability of this must fluctuate both with the skill of the interpreter and the consistency of the judges.

It need hardly be said that in few legal systems are judicial powers confined to authoritative determinations of the fact of violation of the primary rules. Most systems have, after some delay, seen the advantages of further centralization of social pressure; and have partially prohibited the use of physical punishments or violent self help by private individuals. Instead they have supplemented the primary rules of obligation by further secondary rules, specifying or at least limiting the penalties for violation, and have conferred upon judges, where they have ascertained the fact of violation, the exclusive power to direct the application of penalties by other officials. These secondary rules provide the centralized official "sanctions" of the system.

If we stand back and consider the structure which has resulted from the combination of primary rules of obligation with the secondary rules of recognition, change and adjudication, it is plain that we have here not only the heart of a legal system, but a most powerful tool for the analysis of much that has puzzled both the jurist and the political theorist.

RULE OF RECOGNITION AND LEGAL VALIDITY

. . . Wherever . . . a rule of recognition is accepted, both private persons and officials are provided with authoritative criteria for identifying primary rules of obligation. The criteria so provided may, as we have seen, take any one or more of a variety of forms: these include reference to an authoritative text; to legislative enactment; to customary practice; to general declarations of specified persons, or to past judicial decisions in particular cases. In a very simple system like the world of Rex I . . . where only what he enacts is law and no legal limitations upon his legislative power are imposed by customary rule or constitutional document, the sole criterion for identifying the law will be a simple reference to the fact of enactment by Rex I. The existence of this simple form of rule of recognition will be manifest in the general practice, on the part of officials or private persons, of identifying the rules by this criterion. In a modern legal system where there are a variety of "sources" of law, the rule of recognition is correspondingly more complex: the criteria for identifying the law are multiple and commonly include a written constitution, enactment by a legislature, and judicial precedents. In most cases, provision is made for possible conflict by ranking these criteria in an order of relative subordination and primacy. It is in this way that in our system "common law" is subordinate to "statute."

. . . For the most part the rule of recognition is not stated, but its existence is *shown* in the way in which particular rules are identified, either by courts or other officials or private persons or their advisers. There is, of course, a difference in the use made by courts of the criteria provided by the rule and the use of them by others: for when courts reach a particular conclusion on the footing that a particular rule has been correctly identified as law, what they say has a special authoritative status conferred on it by other rules. In this respect, as in many others, the rule of recognition of a legal system is like the scoring rule of a game. In the course of the game the general rule defining the activities which constitute scoring (runs, goals, &c.) is seldom formulated; instead it is *used* by officials and players in identifying the particular phases which count towards winning. Here too, the declarations of officials (umpire or scorer) have a special authoritative status attributed to them by other rules. Further, in both cases there is the possibility of a conflict between these authoritative applications of the rule and the general understanding of what the rule plainly requires according to its terms. This, as we shall see later, is a complication which must be catered for in any account of what it is for a system of rules of this sort to exist.

The use of unstated rules of recognition, by courts and others, in identifying particular rules of the system is characteristic of the internal point of view. Those who use them in this way thereby manifest their own acceptance of them as guiding rules and with this attitude there goes a characteristic vocabulary different from the natural expressions of the external point of view. Perhaps the simplest of these

is the expression, "It is the law that . . .", which we may find on the lips not only of judges, but of ordinary men living under a legal system, when they identify a given rule of the system. This, like the expression "Out" or "Goal," is the language of one assessing a situation by reference to rules which he in common with others acknowledges as appropriate for this purpose. . . . The natural expression of th[e] external point of view is not "It is the law that . . ." but "In England they recognize as law . . . whatever the Queen in Parliament enacts. . . ." The first of these forms of expression we shall call an *internal statement* because it manifests the internal point of view and is naturally used by one who, accepting the rule of recognition and without stating the fact that it is accepted, applies the rule in recognizing some particular rule of the system as valid. The second form of expression we shall call an *external statement* because it is the natural language of an external observer of the system who, without himself accepting its rule of recognition, states the fact that others accept it.

If this use of an accepted rule of recognition in making internal statements is understood and carefully distinguished from an external statement of fact that the rule is accepted, many obscurities concerning the notion of legal "validity" disappear. For the word "valid" is most frequently, though not always, used, in just such internal statements, applying to a particular rule of a legal system, an unstated but accepted rule of recognition. To say that a given rule is valid is to recognize it as passing all the tests provided by the rule of recognition and so as a rule of the system. We can indeed simply say that the statement that a particular rule is valid means that it satisfies all the criteria provided by the rule of recognition. This is incorrect only to the extent that it might obscure the internal character of such statements; for, like the cricketers' "Out" these statements of validity normally apply to a particular case a rule of recognition accepted by the speaker and others, rather than expressly state that the rule is satisfied. . . .

There are, indeed, many questions which we can raise about this ultimate rule. We can ask whether it is the practice of courts, legislatures, officials, or private citizens in England actually to use this rule as an ultimate rule of recognition. Or has our process of legal reasoning been an idle game with the criteria of validity of a system now discarded? We can ask whether it is a satisfactory form of legal system which has such a rule at its root. Does it produce more good than evil? Are there prudential reasons for supporting it? Is there a moral obligation to do so? These are plainly very important questions; but, equally plainly, when we ask them about the rule of recognition, we are no longer attempting to answer the same kind of question about it as those which we answered about other rules with its aid. When we move from saying that a particular enactment is valid, because it satisfies the rule that what the Queen in Parliament enacts is law, to saying that in England this last rule is used by courts, officials, and private persons as the ultimate rule of recognition, we have moved from an internal statement of law asserting the validity of a rule of the system to an external statement of fact which an observer of the system might make even if he did not accept it. So too when we move from the statement that a particular enactment is valid, to the statement that the rule of recognition of the system is an excellent one and the system based on it is one worthy of support, we have moved from a statement of legal validity to a statement of value.

[*Editors' Note: Hart later compares the rule of recognition to the meter bar in Paris, which serves to define the measure of length we conventionally call a meter. Just as it is senseless to ask if that bar really is a meter long, so too is it senseless to ask if the rule of recognition is "valid" legally.*]

REVIEW AND DISCUSSION QUESTIONS

1. What are the three defects of a system of primary rules?
2. How does Hart define a "secondary rule"?
3. How do secondary rules solve the three defects he describes?
4. Hart suggests here (and later in his book makes explicit) that the rule of recognition is neither valid nor invalid, but rather is the test for validity. Explain.

5. Even if the rule of recognition is neither valid nor invalid, it could still be criticized on other grounds. Describe some possible rules of recognition and indicate how each might be assessed.
6. What is the rule of recognition where you live? Describe it.
7. Is it important for Hart's theory that a constitution can include rules governing how it is to be amended? Explain.
8. What are the differences between internal statements and external statements about the law?

The Model of Rules

Ronald Dworkin

In this essay, Ronald Dworkin (who once studied philosophy of law under H. L. A. Hart and once held the same position Hart held at Oxford) offers a careful, detailed criticism of the key positivist claims. Beginning with a brief summary of how both Austin and Hart exemplify the positivist position, he goes on to attack what he sees as positivism's three basic tenets: that law is a set of rules identifiable by their "pedigree," that the law is limited to those rules, and that a person has a legal obligation to do (or not do) something only when a valid legal rule requires the person to do (or not do) it.

I want to examine the soundness of positivism, particularly in the powerful form that Professor H. L. A. Hart of Oxford has given to it. I choose to focus on his position, not only because of its clarity and elegance, but because here, as almost everywhere else in legal philosophy, constructive thought must start with a consideration of his views.

POSITIVISM

Positivism has a few central and organizing propositions as its skeleton, and though not every philosopher who is called a positivist would subscribe to these in the way I present them, they do define the general position I want to examine. These key tenets may be stated as follows:

A. The law of a community is a set of special rules used by the community directly or indirectly for the purpose of determining which behavior will be punished or coerced by the public power. These special rules can be identified and distinguished by specific criteria, by tests having to do not with their content but with their *pedigree* or the manner in which they were adopted or developed. These tests of pedigree can be used to distinguish valid legal rules from spurious legal rules (rules which lawyers and litigants wrongly argue are rules of law) and also from other sorts of social rules (generally lumped together as "moral rules") that the community follows but does not enforce through public power.

B. The set of these valid legal rules is exhaustive of "the law," so that if someone's case is not clearly covered by such a rule (because there is none that seems appropriate, or those that seem appropriate are vague, or for some other reason) then that case cannot be decided by "applying the law." It must be decided by some official, like a judge, "exercising his discretion," which means reaching beyond the law for some other sort of standard to guide him in manufacturing a fresh legal rule or supplementing an old one.

C. To say that someone has a "legal obligation" is to say that his case falls under a valid legal rule that requires him to do or to forbear from doing something. (To say he has a legal right, or has a legal power of some sort, or a legal privilege or immunity, is to assert, in a shorthand way, that others have actual or hypothetical legal obligations to act or not to act in certain ways touching him.) In the absence of such a valid legal rule there is no legal obligation; it follows that when the judge decides an issue by exercising his discretion, he is not enforcing a legal obligation as to that issue.

This is only the skeleton of positivism. The flesh is arranged differently by different positivists, and some even tinker with the bones. Different versions differ chiefly in their description of the fundamental

test of pedigree a rule must meet to count as a rule of law.

Austin, for example, framed his version of the fundamental test as a series of interlocking definitions and distinctions. He defined having an obligation as lying under a rule, a rule as a general command, and a command as an expression of desire that others behave in a particular way, backed by the power and will to enforce that expression in the event of disobedience. He distinguished classes of rules (legal, moral or religious) according to which person or group is the author of the general command the rule represents. In each political community, he thought, one will find a sovereign—a person or a determinate group whom the rest obey habitually, but who is not in the habit of obeying anyone else. The legal rules of a community are the general commands its sovereign has deployed. Austin's definition of legal obligation followed from this definition of law. One has a legal obligation, he thought, if one is among the addressees of some general order of the sovereign, and is in danger of suffering a sanction unless he obeys that order.

Of course, the sovereign cannot provide for all contingencies through any scheme of orders, and some of his orders will inevitably be vague or have furry edges. Therefore (according to Austin) the sovereign grants those who enforce the law (judges) discretion to make fresh orders when novel or troublesome cases are presented. The judges then make new rules or adapt old rules, and the sovereign either overturns their creations, or tacitly confirms them by failing to do so.

Austin's model is quite beautiful in its simplicity. It asserts the first tenet of positivism, that the law is a set of rules specially selected to govern public order, and offers a simple factual test—what has the sovereign commanded?—as the sole criterion for identifying those special rules. In time, however, those who studied and tried to apply Austin's model found it too simple. Many objections were raised, among which were two that seemed fundamental. First, Austin's key assumption that in each community a determinate group or institution can be found, which is in ultimate control of all other groups, seemed not to hold in a complex society. Political control in a modern nation is pluralistic and shifting, a matter of more or less, of compromise and cooperation and alliance, so that it is often impossible to say that any person or group has that dramatic control necessary to qualify as an Austinian sovereign. One wants to say, in the United States for example, that the "people" are sovereign. But this means almost nothing, and in itself provides no test for determining what the "people" have commanded, or distinguishing their legal from their social or moral commands.

Second, critics began to realize that Austin's analysis fails entirely to account for, even to recognize, certain striking facts about the attitudes we take toward "the law." We make an important distinction between law and even the general orders of a gangster. We feel that the law's strictures—and its sanctions—are different in that they are obligatory in a way that the outlaw's commands are not. Austin's analysis has no place for any such distinction, because it defines an obligation as subjection to the threat of force, and so founds the authority of law entirely on the sovereign's ability and will to harm those who disobey. Perhaps the distinction we make is illusory—perhaps our feelings of some special authority attaching to the law is based on religious hangover or another sort of mass self-deception. But Austin does not demonstrate this, and we are entitled to insist that an analysis of our concept of law either acknowledge and explain our attitudes, or show why they are mistaken.

H. L. A. Hart's version of positivism is more complex than Austin's, in two ways. First, he recognizes, as Austin did not, the rules are of different logical kinds (Hart distinguishes two kinds, which he calls "primary" and "secondary" rules). Second, he rejects Austin's theory that a rule is a kind of command, and substitutes a more elaborate general analysis of what rules are. We must pause over each of these points, and then note how they merge in Hart's concept of law.

Hart's distinction between primary and secondary rules is of great importance.[1] Primary rules are those that grant rights or impose obligations upon members of the community. The rules of the criminal law that forbid us to rob, murder or drive too fast are good examples of primary rules. Secondary rules are those that stipulate how, and by whom, such primary rules may be formed, recognized, modified or extinguished. The rules that stipulate how Congress

is composed, and how it enacts legislation, are examples of secondary rules. Rules about forming contracts and executing wills are also secondary rules because they stipulate how very particular rules governing particular legal obligations (that is, the terms of a contract or the provisions of a will) come into existence and are changed.

His general analysis of rules is also of great importance. Austin had said that every rule is a general command, and that a person is obligated under a rule if he is liable to be hurt should be disobey it. Hart points out that this obliterates the distinction between being *obliged* to do something and being *obligated* to do it. If one is bound by a rule he is obligated, not merely obliged, to do what it provides, and therefore being bound by a rule must be different from being subject to an injury if one disobeys an order. A rule differs from an order, among other ways, by being *normative,* by setting a standard of behavior that has a call on its subject beyond the threat that may enforce it. A rule can never be binding just because some person with physical power wants it to be so. He must have *authority* to issue the rule or it is no rule, and such authority can only come from another rule which is already binding on those to whom he speaks. That is the difference between a valid law and the orders of a gunman.

So Hart offers a general theory of rules that does not make their authority depend upon the physical power of their authors. If we examine the way different rules come into being, he tells us, and attend to the distinction between primary and secondary rules, we see that there are two possible sources of a rule's authority.[2]

> **A.** A rule may become binding upon a group of people because that group through its practices *accepts* the rule as a standard for its conduct. It is not enough that the group simply conforms to a pattern of behavior: even though most Englishmen may go to the movies on Saturday evening, they have not accepted a rule requiring that they do so. A practice constitutes the acceptance of a rule only when those who follow the practice regard the rule as binding, and recognize the rule as a reason or justification for their own behavior and as a reason for criticizing the behavior of others who do not obey it.
>
> **B.** A rule may also become binding in quite a different way, namely by being enacted in conformity with some *secondary* rule that stipulates that rules so enacted shall be binding. If the constitution of a club stipulates, for example, that by-laws may be adopted by a majority of the members, then particular by-laws so voted are binding upon all the members, not because of any practice of acceptance of these particular by-laws, but because the constitution says so. We use the concept of *validity* in this connection: rules binding because they have been created in a manner stipulated by some secondary rule are called "valid" rules. Thus we can record Hart's fundamental distinction this way: a rule may be binding (a) because it is accepted or (b) because it is valid.

Hart's concept of law is a construction of these various distinctions. Primitive communities have only primary rules, and these are binding entirely because of practices of acceptance. Such communities cannot be said to have "law," because there is no way to distinguish a set of legal rules from amongst other social rules, as the first tenet of positivism requires. But when a particular community has developed a fundamental secondary rule that stipulates how legal rules are to be identified, the idea of a distinct set of legal rules, and thus of law, is born.

Hart calls such a fundamental secondary rule a "rule of recognition." The rule of recognition of a given community may be relatively simple ("What the king enacts is law") or it may be very complex (the United States Constitution, with all its difficulties of interpretation, may be considered a single rule of recognition). The demonstration that a particular rule is valid may therefore require tracing a complicated chain of validity back from that particular rule ultimately to the fundamental rule. Thus a parking ordinance of the city of New Haven is valid because it is adopted by a city council, pursuant to the procedures and within the competence specified by the municipal law adopted by the state of Connecticut, in conformity with the procedures and within the competence specified by the constitution of the state of Connecticut, which was in turn adopted consistently with the requirements of the United States Constitution.

Of course, a rule of recognition cannot itself be valid, because by hypothesis it is ultimate, and so cannot meet tests stipulated by a more fundamental rule. The rule of recognition is the sole rule in a legal system whose binding force depends upon its

acceptance. If we wish to know what rule of recognition a particular community has adopted or follows, we must observe how its citizens, and particularly its officials, behave. We must observe what ultimate arguments they accept as showing the validity of a particular rule, and what ultimate arguments they use to criticize other officials or institutions. We can apply no mechanical test, but there is no danger of our confusing the rule of recognition of a community with its rules of morality. The rule of recognition is identified by the fact that its province is the operation of the governmental apparatus of legislatures, courts, agencies, policemen, and the rest.

In this way Hart rescues the fundamentals of positivism from Austin's mistakes. Hart agrees with Austin that valid rules of law may be created through the acts of officials and public institutions. But Austin thought that the authority of these institutions lay only in their monopoly of power. Hart finds their authority in the background of constitutional standards against which they act, constitutional standards that have been accepted, in the form of a fundamental rule of recognition, by the community which they govern. This background legitimates the decisions of government and gives them the cast and call of obligation that the naked commands of Austin's sovereign lacked. Hart's theory differs from Austin's also, in recognizing that different communities use different ultimate tests of law, and that some allow other means of creating law than the deliberate act of a legislative institution. Hart mentions "long customary practice" and "the relation [of a rule] to judicial decisions" as other criteria that are often used, though generally along with and subordinate to the test of legislation.

So Hart's version of positivism is more complex than Austin's, and his test for valid rules of law is more sophisticated. In one respect, however, the two models are very similar. Hart, like Austin, recognizes that legal rules have fuzzy edges (he speaks of them as having "open texture") and, again like Austin, he accounts for troublesome cases by saying that judges have had to exercise discretion to decide these cases by fresh legislation. (I shall later try to show why one who thinks of law as a special set of rules is almost inevitably drawn to account for difficult cases in terms of someone's exercise of discretion.)

RULES, PRINCIPLES, AND POLICIES

I want to make a general attack on positivism, and I shall use H. L. A. Hart's version as a target, when a particular target is needed. My strategy will be organized around the fact that when lawyers reason or dispute about legal rights and obligations, particularly in those hard cases when our problems with these concepts seem most acute, they make use of standards that do not function as rules, but operate differently as principles, policies, and other sorts of standards. Positivism, I shall argue, is a model of and for a system of rules, and its central notion of a single fundamental test for law forces us to miss the important roles of these standards that are not rules.

I just spoke of "principles, policies, and other sorts of standards." Most often I shall use the term "principle" generically, to refer to the whole set of these standards other than rules; occasionally, however, I shall be more precise, and distinguish between principles and policies. Although nothing in the present argument will turn on the distinction, I should state how I draw it. I call a "policy" that kind of standard that sets out a goal to be reached, generally an improvement in some economic, political, or social feature of the community (though some goals are negative, in that they stipulate that some present feature is to be protected from adverse change). I call a "principle" a standard that is to be observed, not because it will advance or secure an economic, political, or social situation deemed desirable, but because it is a requirement of justice or fairness or some other dimension of morality. Thus the standard that automobile accidents are to be decreased is a policy, and the standard that no man may profit by his own wrong a principle. . . .

My immediate purpose, however, is to distinguish principles in the generic sense from rules, and I shall start by collecting some examples of the former. The examples I offer are chosen haphazardly; almost any case in a law school casebook would provide examples that would serve as well. In 1889 a New York court, in the famous case of *Riggs v. Palmer* [reprinted in this book—EDS.] had to decide whether an heir named in the will of his grandfather could inherit under that will, even though he had murdered his grandfather to do so.

The court began its reasoning with this admission: "It is quite true that statutes regulating the making, proof and effect of wills, and the devolution of property, if literally construed, and if their force and effect can in no way and under no circumstances be controlled or modified, give this property to the murderer." But the court continued to note that "all laws as well as all contracts may be controlled in their operation and effect by general, fundamental maxims of the common law. No one shall be permitted to profit by his own fraud, or to take advantage of his own wrong, or to found any claim upon his own iniquity, or to acquire property by his own crime." The murderer did not receive his inheritance.

In 1960, a New Jersey court was faced, in *Henningsen v. Bloomfield Motors, Inc.,*[3] with the important question of whether (or how much) an automobile manufacturer may limit his liability in case the automobile is defective. Henningsen had bought a car, and signed a contract which said that the manufacturer's liability for defects was limited to "making good" defective parts—"this warranty being expressly in lieu of all other warranties, obligations or liabilities." Henningsen argued that, at least in the circumstances of his case, the manufacturer ought not to be protected by this limitation, and ought to be liable for the medical and other expenses of persons injured in a crash. He was not able to point to any statute, or to any established rule of law, that prevented the manufacturer from standing on the contract. The court nevertheless agreed with Henningsen. At various points in the court's argument the following appeals to standards are made: (a) "[W]e must keep in mind the general principle that, in the absence of fraud, one who does not choose to read a contract before signing it cannot later relieve himself of its burdens." (b) "In applying that principle, the basic tenet of freedom of competent parties to contract is a factor of importance." (c) "Freedom of contract is not such an immutable doctrine as to admit of no qualification in the area in which we are concerned." (d) "In a society such as ours where the automobile is a common and necessary adjunct of daily life, and where its use is so fraught with danger to the driver, passengers and the public, the manufacturer is under a special obligation in connection with the construction, promotion and sale of his cars. Consequently, the courts must examine purchase agreements closely to see if consumer and public interests are treated fairly." (e) "[I]s there any principle which is more familiar or more firmly embedded in the history of Anglo-American law than the basic doctrine that the courts will not permit themselves to be used as instruments of inequity and injustice?"[4] (f) "More specifically, the courts generally refuse to lend themselves to the enforcement of a 'bargain' in which one party has unjustly taken advantage of the economic necessities of [the] other. . . ."[5]

The standards set out in these quotations are not the sort we think of as legal rules. They seem very different from propositions like "The maximum legal speed on the turnpike is sixty miles an hour" or "A will is invalid unless signed by three witnesses." They are different because they are legal principles rather than legal rules.

The difference between legal principles and legal rules is a logical distinction. Both sets of standards point to particular decisions about legal obligation in particular circumstances, but they differ in the character of the direction they give. Rules are applicable in an all-or-nothing fashion. If the facts a rule stipulates are given, then either the rule is valid, in which case the answer it supplies must be accepted, or it is not, in which case it contributes nothing to the decision.

This all-or-nothing is seen most plainly if we look at the way rules operate, not in law, but in some enterprise they dominate—a game, for example. In baseball a rule provides that if the batter has had three strikes, he is out. An official cannot consistently acknowledge that this is an accurate statement of a baseball rule, and decide that a batter who has had three strikes is not out. Of course, a rule may have exceptions (the batter who has taken three strikes is not out if the catcher drops the third strike). However, an accurate statement of the rule would take this exception into account, and any that did not would be incomplete. If the list of exceptions is very large, it would be too clumsy to repeat them each time the rule is cited; there is, however, no reason in theory why they could not all be added on, and the more that are, the more accurate is the statement of the rule.

If we take baseball rules as a model, we find that rules of law, like the rule that a will is invalid unless

signed by three witnesses, fit the model well. If the requirement of three witnesses is a valid legal rule, then it cannot be that a will has been signed by only two witnesses and is valid. The rule might have exceptions, but if it does then it is inaccurate and incomplete to state the rule so simply, without enumerating the exceptions. In theory, at least, the exceptions could all be listed, and the more of them that are, the more complete is the statement of the rule.

But this is not the way the sample principles in the quotations operate. Even those which look most like rules do not set out legal consequences that follow automatically when the conditions provided are met. We say that our law respects the principle that no man may profit from his own wrong, but we do not mean that the law never permits a man to profit from wrongs he commits. In fact, people often profit, perfectly legally, from their legal wrongs. The most notorious case is adverse possession—if I trespass on your land long enough, some day I will gain a right to cross your land whenever I please. There are many less dramatic examples. If a man leaves one job, breaking a contract, to take a much higher paying job, he may have to pay damages to his first employer, but he is usually entitled to keep his new salary. If a man jumps bail and crosses state lines to make a brilliant investment in another state, he may be sent back to jail, but he will keep his profits.

We do not treat these—and countless other counter-instances that can easily be imagined—as showing that the principle about profiting from one's wrongs is not a principle of our legal system, or that it is incomplete and needs qualifying exceptions. We do not treat counter-instances as exceptions (at least not exceptions in the way in which a catcher's dropping the third strike is an exception) because we could not hope to capture these counter-instances simply by a more extended statement of the principle. They are not, even in theory, subject to enumeration, because we would have to include not only these cases (like adverse possession) in which some institution has already provided that profit can be gained through a wrong, but also those numberless imaginary cases in which we know in advance that the principle would not hold. Listing some of these might sharpen our sense of the principle's weight (I shall mention that dimension in a moment), but it would not make for a more accurate or complete statement of the principle.

A principle like "No man may profit from his own wrong" does not even purport to set out conditions that make its application necessary. Rather, it states a reason that argues in one direction, but does not necessitate a particular decision. If a man has or is about to receive something, as a direct result of something illegal he did to get it, then that is a reason which the law will taken into account in deciding whether he should keep it. There may be other principles or policies arguing in the other direction—a policy of securing title, for example, or a principle limiting punishment to what the legislature has stipulated. If so, our principle may not prevail, but that does not mean that it is not a principle of our legal system, because in the next case, when these contravening considerations are absent or less weighty, the principle may be decisive. All that is meant, when we say that a particular principle is a principle of our law, is that the principle is one which officials must take into account, if it is relevant, as a consideration inclining in one direction or another.

The logical distinction between rules and principles appears more clearly when we consider principles that do not even look like rules. Consider the proposition, set out under "(d)" in the excerpts from the *Henningsen* opinion, that "the manufacturer is under a special obligation in connection with the construction, promotion and sale of his cars." This does not even purport to define the specific duties such a special obligation entails, or to tell us what rights automobile consumers acquire as a result. It merely states—and this is an essential link in the *Henningsen* argument—that automobile manufacturers must be held to higher standards than other manufacturers, and are less entitled to rely on the competing principle of freedom of contract. It does not mean that they may never rely on that principle, or that courts may rewrite automobile purchase contracts at will; it means only that if a particular clause seems unfair or burdensome, courts have less reason to enforce the clause than if it were for the purchase of neckties. The "special obligation" counts in favor, but does not in itself necessitate, a decision refusing to enforce the terms of an automobile purchase contract.

This first difference between rules and principles entails another. Principles have a dimension that rules do not—the dimension of weight or importance. When principles intersect (the policy of protecting automobile consumers intersecting with principles of freedom of contract, for example), one who must resolve the conflict has to take into account the relative weight of each. This cannot be, of course, an exact measurement, and the judgment that a particular principle or policy is more important than another will often be a controversial one. Nevertheless, it is an integral part of the concept of a principle that it has this dimension, that it makes sense to ask how important or how weighty it is.

Rules do not have this dimension. We can speak of rules as being *functionally* important or unimportant (the baseball rule that three strikes are out is more important than the rule that runners may advance on a balk, because the game would be much more changed with the first rule altered than the second). In this sense, one legal rule may be more important than another because it has a greater or more important role in regulating behavior. But we cannot say that one rule is more important than another within the system of rules, so that when two rules conflict one supersedes the other by virtue of its greater weight. If two rules conflict, one of them cannot be a valid rule. The decision as to which is valid, and which must be abandoned or recast, must be made by appealing to considerations beyond the rules themselves. A legal system might regulate such conflicts by other rules, which prefer the rule enacted by the higher authority, or the rule enacted later, or the more specific rule, or something of that sort. A legal system may also prefer the rule supported by the more important principles. (Our own legal system uses both of these techniques.) . . .

Sometimes a rule and a principle can play much the same role, and the difference between them is almost a matter of form alone. The first section of the Sherman Act states that every contract in restraint of trade shall be void. The Supreme Court had to make the decision whether this provision should be treated as a rule in its own terms (striking down every contract "which restrains trade," which almost any contract does) or as a principle, providing a reason for striking down a contract in the absence of effective contrary policies. The Court construed the provision as a rule, but treated that rule as containing the word "unreasonable," and as prohibiting only "unreasonable" restraints of trade.[6] This allowed the provision to function logically as a rule (whenever a court finds that the restraint is "unreasonable" it is bound to hold the contract invalid) and substantially as a principle (a court must take into account a variety of other principles and policies in determining whether a particular restraint in particular economic circumstances is "unreasonable").

Words like "reasonable," "negligent," "unjust," and "significant" often perform just this function. Each of these terms makes the application of the rule which contains it depend to some extent upon principles or policies lying beyond the rule, and in this yaw makes that rule itself more like a principle. . . .

Once we identify legal principles as separate sorts of standards, different from legal rules, we are suddenly aware of them all around us. Law teachers teach them, lawbooks cite them, legal historians celebrate them. But they seem most energetically at work, carrying most weight, in difficult lawsuits like *Riggs* and *Henningsen.* In cases like these principles play an essential part in arguments supporting judgments about particular legal rights and obligations. . . .

THE RULE OF RECOGNITION

. . . If principles of the *Riggs* and *Henningsen* sort are to count as law, and we are nevertheless to preserve the notion of a master rule for law, then we must be able to deploy some test that all (and only) the principles that do count as law meet. Let us begin with the test Hart suggests for identifying valid *rules* of law, to see whether these can be made to work for principles as well.

Most rules of law, according to Hart, are valid because some competent institution enacted them. Some were created by a legislature, in the form of statutory enactments. Others were created by judges who formulated them to decide particular cases, and thus established them as precedents for the future. But this test of pedigree will not work for the *Riggs* and *Henningsen* principles. The origin of these as legal principles lies not in a particular decision of

some legislature or court, but in a sense of appropriateness developed in the profession and the public over time. Their continued power depends upon this sense of appropriateness being sustained. If it no longer seemed unfair to allow people to profit by their wrongs, or fair to place special burdens upon oligopolies that manufacture potentially dangerous machines, these principles would no longer play much of a role in new cases, even if they had never been overruled or repealed. (Indeed, it hardly makes sense to speak of principles like these as being "overruled" or "repealed." When they decline they are eroded, not torpedoed.)

True, if we were challenged to back up our claim that some principle is a principle of law, we would mention any prior cases in which that principle was cited, or figured in the argument. We would also mention any statute that seemed to exemplify that principle (even better if the principle was cited in the preamble of the statute, or in the committee reports or other legislative documents that accompanied it). Unless we could find some such institutional support, we would probably fail to make out our case, and the more support we found, the more weight we could claim for the principle.

Yet we could not devise any formula for testing how much and what kind of institutional support is necessary to make a principle a legal principle, still less to fix its weight at a particular order of magnitude. We argue for a particular principle by grappling with a whole set of shifting, developing and interacting standards (themselves principles rather than rules) about institutional responsibility, statutory interpretation, the persuasive force of various sorts of precedent, the relation of all these to contemporary moral practices, and hosts of other such standards. We could not bolt all of these together into a single "rule," even a complex one, and if we could the result would bear little relation to Hart's picture of a rule of recognition, which is the picture of a fairly stable master rule specifying "some feature or features possession of which by a suggested rule is taken as a conclusive affirmative indicating that it is a rule. . . ."

If we are asked (as we might well be) to defend the particular doctrine of precedent, or the particular technique of statutory interpretation, that we used in this argument, we should certainly cite the practice of others in using that doctrine or technique. But we should also cite other general principles that we believe support that practice, and this introduces a note of validity into the chord of acceptance. We might argue, for example, that the use we make of earlier cases and statutes is supported by a particular analysis of the point of practice of legislation or the doctrine of precedent, or by the principles of democratic theory, or by a particular position on the proper division of authority between national and local institutions, or something else of that sort. Nor is this path of support a one-way street leading to some ultimate principle resting on acceptance alone. Our principles of legislation, precedent, democracy, or federalism might be challenged too; and if they were we should argue for them, not only in terms of practice, but in terms of each other and in terms of the implications of trends of judicial and legislative decisions, even though this last would involve appealing to those same doctrines of interpretation we justified through the principles we are now trying to support. At this level of abstraction, in other words, principles rather hang together than link together.

So even though principles draw support from the official acts of legal institutions, they do not have a simple or direct enough connection with these acts to frame that connection in terms of criteria specified by some ultimate master rule of recognition. . . .

DISCRETION

Discretion, like the hole in a doughnut, does not exist except as an area left open by a surrounding belt of restriction. It is therefore a relative concept. It always makes sense to ask, "Discretion under which standards?" or "Discretion as to which authority?" Generally the context will make the answer to this plain, but in some cases the official may have discretion from one stand-point though not from another. . . .

Sometimes we use "discretion" in a weak sense, simply to say that for some reason the standards an official must apply cannot be applied mechanically but demand the use of judgment. We use this weak sense when the context does not already make that clear, when the background our audience assumes

does not contain that piece of information. Thus we might say, The sergeant's orders left him a great deal of "discretion," to those who do not know what the sergeant's orders were or who do not know something that made those orders vague or hard to carry out. It would make perfect sense to add, by way of amplification, that the lieutenant had ordered the sergeant to take his five most experienced men on patrol but that it was hard to determine which were the most experienced.

Sometimes we use the term in a different weak sense, to say only that some official has final authority to make a decision and cannot be reviewed and reversed by any other official. We speak this way when the official is part of a hierarchy of officials structured so that some have higher authority but in which the patterns of authority are different for different classes of decision. Thus we might say that in baseball certain decisions, like the decision whether the ball or the runner reached second base first, are left to the discretion of the second base umpire, if we mean that on this issue the head umpire has no power to substitute his own judgment if he disagrees.

I call both of these senses weak to distinguish them from a stronger sense. We use "discretion" sometimes not merely to say that an official must use judgment in applying the standards set him by authority, or that no one will review that exercise of judgment, but to say that on some issue he is simply not bound by standards set by the authority in question. In this sense we say that a sergeant has discretion who has been told to pick any five men for patrol he chooses or that a judge in a dog show has discretion to judge airedales before boxers if the rules do not stipulate an order of events. We use this sense not to comment on the vagueness or difficulty of the standards, or on who has the final word in applying them, but on their range and the decisions they purport to control. If the sergeant is told to take the five most experienced men, he does not have discretion in this strong sense because that order purports to govern his decision. The boxing referee who must decide which fighter has been the more aggressive does not have discretion, in the strong sense, for the same reason.

If anyone said that the sergeant or the referee had discretion in these cases, we should have to understand him, if the context permitted, as using the term in one of the weak senses. Suppose, for example, the lieutenant ordered the sergeant to select the five men he deemed most experienced, and then added that the sergeant had discretion to choose them. Or the rules provided that the referee should award the round to the more aggressive fighter, with discretion in selecting him. We should have to understand these statements in the second weak sense, as speaking to the question of review of the decision. The first weak sense—that the decisions take judgment—would be otiose, and the third, strong sense is excluded by the statements themselves. . . .

So we must examine the doctrine of judicial discretion in the strong sense. (I shall henceforth use the term "discretion" in that sense.) Do the principles judges cited in cases like *Riggs* or *Henningsen* control their decisions, as the sergeant's orders to take the most experienced men or the referee's duty to choose the more aggressive fighter control the decisions of these officials? What arguments could a positivist supply to show that they do not?

(1) A positivist might argue that principles cannot be binding or obligatory. That would be a mistake. It is always a question, of course, whether any particular principle is *in fact* binding upon some legal official. But there is nothing in the logical character of a principle that renders it incapable of binding him. Suppose that the judge in *Henningsen* had failed to take any account of the principle that automobile manufacturers have a special obligation to their consumers, or the principle that the courts seek to protect those whose bargaining position is weak, but had simply decided for the defendant by citing the principle of freedom of contract without more. His critics would not have been content to point out that he had not taken account of considerations that other judges have been attending to for some time. Most would have said that it was his duty to take the measure of these principles and that the plaintiff was entitled to have him do so. We mean no more, when we say that a *rule* is binding upon a judge, than that he must follow it if it applies, and that if he does not he will on that account have made a mistake. . . .

(2) A positivist might argue that even though some principles are binding, in the sense that the judge must take them into account, they cannot

determine a particular result. This is a harder argument to assess because it is not clear what it means for a standard to "determine" a result. Perhaps it means that the standard *dictates* the result whenever it applies so that nothing else counts. If so, then it is certainly true that the individual principles do not determine results, but that is only another way of saying that principles are not rules. Only rules dictate results, come what may. When a contrary result has been reached, the rule has been abandoned or changed. Principles do not work that way; they incline a decision one way, though not conclusively, and they survive intact when they do not prevail. . . . If a judge believes that principles he is bound to recognize point in one direction and that principles pointing in the other direction, if any, are not of equal weight, then he must decide accordingly, just as he must follow what he believes to be a binding rule. He may, of course, be wrong in his assessment of the principles, but he may also be wrong in his judgment that the rule is binding. The sergeant and the referee, we might add, are often in the same boat. No one factor dictates which soldiers are the most experienced or which fighter the more aggressive. These officials must make judgments of the relative weights of these various factors; they do not on that account have discretion.

(3) A positivist might argue that principles cannot count as law because their authority, and even more so their weight, are congenitally *controversial*. It is true that generally we cannot *demonstrate* the authority or weight of a particular principle as we can sometimes demonstrate the validity of a rule by locating it in an act of Congress or in the opinion of an authoritative court. Instead, we make a case for a principle, and for its weight, by appealing to an amalgam of practice and other principles in which the implications of legislative and judicial history figure along with appeals to community practices and understandings. There is no litmus paper for testing the soundness of such a case—it is a matter of judgment, and reasonable men may disagree. But again this does not distinguish the judge from other officials who do not have discretion. The sergeant has no litmus paper for experience, the referee none for aggressiveness. Neither of these has discretion, because he is bound to reach an understanding, controversial or not, of what his orders or the rules require, and to act on that understanding. That is the judge's duty as well.

NOTES

1. *See* H. L. A. Hart, *The Concept of Law* 89–96 (1961).
2. *Id.* at 97–107.
3. *32* N.J. 358, 161 A.2d 69 (1960).
4. *Id.* at 389, 161 A.2d at 86 (quoting Frankfurter, J., in *United States v. Bethlehem Steel,* 315 U.S. 289, 326 (1942).
5. *Id.*
6. *Standard Oil v. United States,* 221 U.S. 1, 60 (1911); *United States v. American Tobacco Co.,* 221 U.S. 106, 180 (1911).

REVIEW AND DISCUSSION QUESTIONS

1. What are the three key tenets of positivism, according to Dworkin?
2. Explain the role of principles in *Riggs v. Palmer.*
3. Principles in the general sense include two sorts of standards. What are the two?
4. How do principles differ from rules, according to Dworkin?
5. Why cannot Hart modify his view to include principles, according to Dworkin?
6. What is the distinction between "weak" and "strong" discretion?
7. How does Dworkin answer positivists who claim judges exercise "discretion" in the strong sense?

Natural Law Revisited

Ronald Dworkin

Ronald Dworkin begins this essay with a brief discussion of "naturalism," or natural law, indicating that the view he wishes to defend can best be understood to fall within that general camp. He goes on to explain his theory of law with the example of a "chain novel" in which each author is responsible for taking the chapters by previous writers and continuing the novel. With this he illustrates the two dimensions of legal or literary interpretation: fit and justification. He then explains how his theory of law, which began with questions of interpretation, belongs within the natural law tradition.

Everyone likes categories, and legal philosophers like them very much. So we spend a good deal of time, not all of it profitably, labeling ourselves and the theories of law we defend. One label, however, is particularly dreaded: no one wants to be called a natural lawyer. Natural law insists that what the law is depends in some way on what the law should be. This seems metaphysical or at least vaguely religious. In any case it seems plainly wrong. If some theory of law is shown to be a natural law theory, therefore, people can be excused if they do not attend to it much further.

. . . If the crude description of natural law I just gave is correct, that any theory which makes the content of law sometimes depend on the correct answer to some moral question is a natural law theory, then I am guilty of natural law. I am not now interested, I should add, in whether this crude characterization is historically correct, or whether it succeeds in distinguishing natural law from positivist theories of law. My present concern is rather this. Suppose this *is* natural law. What in the world is wrong with it?

A. NATURALISM

I shall start by giving the picture of adjudication I want to defend a name, and it is a name which accepts the crude characterization. I shall call this picture naturalism. According to naturalism, judges should decide hard cases by interpreting the political structure of their community in the following, perhaps special way, by trying to find the best *justification* they can find, in principles of political morality, for the structure as a whole, from the most profound constitutional rules and arrangements to the details of, for example, the private law of tort or contract. Suppose the question arises for the first time, for example, whether and in what circumstances careless drivers are liable, not only for physical injuries to those whom they run down, but also for any emotional damage suffered by relatives of the victim who are watching. According to naturalism, judges should then ask the following questions of the history (including the contemporary history) of their political structure. Does the best possible justification of that history suppose a principle according to which people who are injured emotionally in this way have a right to recover damages in court? If so, what, more precisely, is that principle? Does it entail, for example, that only immediate relatives of the person physically injured have that right? Or only relatives on the scene of the accident, who might themselves have suffered physical damage?

Of course a judge who is faced with these questions in an actual case cannot undertake anything like a full justification of all parts of the constitutional arrangement, statutory system and judicial precedents that make up his "law." I had to invent a mythical judge, called Hercules, with superhuman powers in order even to contemplate what a full justification of the entire system would be like.[1] Real judges can attempt only what we might call a partial justification of the law. They can try to justify, under some set of principles, those parts of the legal background which seem to them immediately relevant, like, for example, the prior judicial decisions about recovery for various sorts of damage in automobile accidents. Nevertheless it is useful to describe this as a partial justification—as a part of what Hercules

himself would do—in order to emphasize that, according to this picture, a judge should regard the law he mines and studies as embedded in a much larger system, so that it is always relevant for him to expand his investigation by asking whether the conclusions he reaches are consistent with what he would have discovered had his study been wider.

It is obvious why this theory of adjudication invites the charge of natural law. It makes each judge's decision about the burden of past law depend on his judgment about the best political justification of that law, and this is of course a matter of political morality. Before I consider whether this provides a fatal defect in the theory, however, I must try to show how the theory might work in practice. . . .

B. THE CHAIN NOVEL

Imagine, then, that a group of novelists is engaged for a particular project. They draw lots to determine the order of play. The lowest number writes the opening chapter of a novel, which he then sends to the next number, who is given the following assignment. He must add a chapter to that novel, which he must write so as to make the novel being constructed the best novel it can be. When he completes his chapter, he then sends the two chapters to the next novelist, who has the same assignment, and so forth. Now every novelist but the first has the responsibility of interpreting what has gone before in the sense of interpretation I described for a naturalist judge. Each novelist must decide what the characters are "really" like; what motives in fact guide them; what the point or theme of the developing novel is; how far some literary device or figure consciously or unconsciously used can be said to contribute to these, and therefore should be extended, refined, trimmed or dropped. He must decide all this in order to send the novel further in one direction rather than another. But all these decisions must be made, in accordance with the directions given, by asking which decisions make the continuing novel better as a novel.

Some novels have in fact been written in this way (including the soft-core pornographic novel *Naked Came the Stranger*) though for a debunking purpose, and certain parlor games, for rainy weekends in English country houses, have something of the same structure. But in this case the novelists are expected to take their responsibilities seriously, and to recognize the duty to create, so far as they can, a single unified novel rather than, for example, a series of independent short stories with characters bearing the same names. Perhaps this is an impossible assignment; perhaps the project is doomed to produce, not simply an impossibly bad novel, but no novel at all, because the best theory of art requires a single creator, or if more than one, that each have some control over the whole. (But what about legends and jokes? What about the Old Testament, or, on some theories, the *Iliad?*) I need not push that question further, because I am interested only in the fact that the assignment makes sense, that each of the novelists in the chain can have some sense of what he or she is asked to do, whatever misgivings each might have about the value or character of what will then be produced.

The crucial question each must face is this. What is the difference between continuing the novel in the best possible way, by writing plot and development that can be seen to flow from what has gone before, and starting a fresh novel with characters having the same names? Suppose you are a novelist well down the chain, and are handed several chapters which are, in fact, the first sections of the Dickens short novel *A Christmas Carol.* You consider these two interpretations of the central character: that Scrooge is irredeemably, inherently evil, and so an example of the degradation of which human nature is intrinsically capable, or that Scrooge is inherently good, but progressively corrupted by the false values and perverse demands of high capitalist society. The interpretation you adopt will obviously make an enormous difference in the way you continue the story. You aim, in accordance with your instructions, to make the continuing novel the best novel it can be; but you must nevertheless choose an interpretation that makes the novel a single work of art. So you will have to respect the text you have been given, and not choose an interpretation that you believe the text rules out. The picture that text gives of Scrooge's early life, for example, might be incompatible with the claim that he is inherently wicked. In that case you have no choice. If, on the other hand, the text is equally consistent with both

interpretations, then you do have a choice. You will choose the interpretation that you believe makes the work more significant or otherwise better, and this will probably (though not inevitably) depend on whether you think people like Scrooge are in fact, in the real world, born bad or corrupted by capitalism.

Now consider a more complex case. Suppose the text does not absolutely rule out either interpretation, but is marginally less consistent with one, which is, however, the interpretation you would pick if they both fit equally well. Suppose you believe that the original sin interpretation (as we might call it) is much the more accurate depiction of human nature. But if you choose that interpretation you will have to regard certain incidents and attributions established in the text you were given as "mistakes." You must then ask yourself which interpretation makes the work of art better *on the whole,* recognizing, as you will, that a novel whose plot is inconsistent or otherwise lacks integrity is thereby flawed. You must ask whether the novel is still better as a novel, read as a study of original sin, even though it must now be regarded as containing some "mistakes" in plot, than it would be with fewer "mistakes" but a less revealing picture of human nature. You may never have reflected on that question before, but that is no reason why you may not do so now, and once you make up your mind you will believe that the correct interpretation of Scrooge's character is the interpretation that makes the novel better on the whole.

C. THE CHAIN OF LAW

Naturalism is a theory of adjudication, not of the interpretation of novels. But naturalism supposes that common law adjudication is a chain enterprise sharing many of the features of the story we invented. According to naturalism, a judge should decide fresh cases in the spirit of a novelist in the chain writing a fresh chapter. The judge must make creative decisions, but must try to make these decisions "going on as before" rather than by starting in a new direction as if writing on a clean slate. He must read through (or have some good idea through his legal training and experience) what other judges in the past have written, not simply to discover what these other judges have said, or their state of mind when they said it, but to reach an opinion about what they have collectively *done,* in the way that each of our novelists formed an opinion about the collective novel so far written. Of course, the best interpretation of past judicial decisions is the interpretation that shows these in the best light, not aesthetically but politically, as coming as close to the correct ideals of a just legal system as possible. Judges in the chain of law share with the chain novelists the imperative of interpretation, but they bring different standards of success—political rather than aesthetic—to bear on that enterprise.

The analogy shows, I hope, how far naturalism allows a judge's beliefs about the personal and political rights people have "naturally"—that is, apart from the law—to enter his judgments about what the law requires. It does not instruct him to regard these beliefs as the only test of law. A judge's background and moral convictions will influence his decisions about what legal rights people have under the law. But the brute facts of legal history will nevertheless limit the role these convictions can play in those decisions. The same distinction we found in literary interpretation, between interpretation and ideal, holds here as well. An Agatha Christie mystery thriller cannot be interpreted as a philosophical novel about the meaning of death even by someone who believes that a successful philosophical novel would be a greater literary achievement than a successful mystery. It cannot be interpreted that way because, if it is, too much of the book must be seen as accidental, and too little as integrated, in plot, style and trope, with its alleged genre or point. Interpreted that way it becomes a shambles and so a failure rather than a success at anything at all. In the same way, a judge cannot plausibly discover, in a long and unbroken string of prior judicial decisions in favor of the manufacturers of defective products, any principle establishing strong consumers' rights. For that discovery would not show the history of judicial practice in a better light; on the contrary it would show it as the history of cynicism and inconsistency, perhaps of incoherence. A naturalist judge must show the facts of history in the best light he can, and this means that he must not show that history as unprincipled chaos.

Of course this responsibility, for judges as well as novelists, may best be fulfilled by a dramatic reinterpretation that both unifies what has gone before and gives it new meaning or point. This explains why a naturalist decision, though it is in this way tied to the past, may yet seem radical. A naturalist judge might find, in some principle that has not yet been recognized in judicial argument, a brilliantly unifying account of past decisions that shows them in a better light than ever before. American legal education celebrates dozens of such events in our own history. In the most famous single common law decision in American jurisprudence, for example, Cardozo reinterpreted a variety of cases to find, in these cases, the principle on which the modern law of negligence was built.

Nevertheless the constraint, that a judge must continue the past and not invent a better past, will often have the consequence that a naturalist judge cannot reach decisions that he would otherwise, given his own political theory, want to reach. A judge who, as a matter of political conviction, believes in consumers' rights may nevertheless have to concede that the law of his jurisdiction has rejected this idea. It is in one way misleading to say, however, that he will be then forced to make decisions *at variance with* his political convictions. The principle that judges should decide consistently with principle, and that law should be coherent, is part of his convictions, and it is this principle that makes the decision he otherwise opposes necessary.

D. INTERPRETATION IN PRACTICE

In this section I shall try to show how a self-conscious naturalist judge might construct a working approach to adjudication, and the role his background moral and political convictions would play in that working approach. When we imagined you to be a novelist in the chain novel, several pages ago, we considered how you would continue the first few chapters of *A Christmas Carol.* We distinguished two dimensions of a successful interpretation. An interpretation must "fit" the data it interprets, in order not to show the novel as sloppy or incoherent, and it must also show that data in its best light, as serving as well as can be some proper ambition of novels. Just now, in noticing how a naturalist judge who believed in consumers' rights might nevertheless have to abandon the claim that consumers' rights are embedded in legal history, . . . the same distinction between these two dimensions was relied upon. A naturalist judge would be forced to reject a politically attractive interpretation, we supposed, simply because he did not believe it fit the record well enough. If fit is indeed an independent dimension of success in interpretation, then any judge's working approach would include some tacit conception of what "fit" is, and of how well a particular interpretation must fit the record of judicial and other legal decisions in order to count as acceptable.

This helps us to explain why two naturalist judges might reach different interpretations of past judicial decisions about accidents, for example. They might hold different conceptions of "fit" or "best fit," so that, for instance, one thinks that an interpretation provides an acceptable fit only if it is supported by the opinions of judges in prior cases, while the other thinks it is sufficient, to satisfy the dimension of fit, that an interpretation fit the actual decisions these judges reached even if it finds no echo in their opinions. This difference might be enough to explain, for example, why one judge could accept an "economic" interpretation of the accident cases—that the point of negligence law is to reduce the overall social costs of accidents—while another judge, who also found that interpretation politically congenial, would feel bound by his beliefs about the requirement of fit to reject it.

At some point, however, this explanation of differences between two judges' theories of the same body of law would become strained and artificial. Suppose Judge X believes, for example, that pedestrians ought to look out for themselves, and have no business walking in areas in which drivers are known normally to exceed the legal speed limit. He might rely on this opinion in deciding that "our law recognizes no general right to recover whenever someone is injured by a speeding driver while walking on a highway where most drivers speed." If Judge Y reaches a different judgment about what the law is, because he believes that pedestrians should be entitled to assume that people will obey the law even when there is good evidence that they will not, then it

would strain language to explain this difference by saying that these judges disagree about the way or the degree in which an interpretation of the law must fit past decisions. We would do better to say that these judges interpret the law differently, in this instance, because they bring different background theories of political morality to their interpretations just as two art critics might disagree about the correct interpretation of impressionism because they bring different theories about the value of art to that exercise.

Any naturalist judge's working approach to interpretation will recognize this distinction between two "dimensions" of interpretations of the prior law, and so we might think of such a theory as falling into two parts. One part refines and develops the idea that an interpretation must fit the data it interprets. This part takes up positions on questions like the following. How many decisions (roughly) can an interpretation set aside as mistakes, and still count as an interpretation of the string of decisions that includes those "mistakes"? How far is an interpretation better if it is more consistent with later rather than earlier past decisions? How far and in what way must a good interpretation fit the opinions judges write as well as the decisions they make? How far must it take account of popular morality contemporary with the decisions it offers to interpret? A second part of any judge's tacit theory of interpretation, however, will be quite independent of these "formal" issues. It will contain the substantive ideals of political morality on which he relies in deciding whether any putative interpretation is to be preferred because it shows legal practice to be better as a matter of substantive justice. Of course, if any working approach to interpretation has these two parts, then it must also have principles that combine or adjudicate between them.

This account of the main structure of a working theory of interpretation has heuristic appeal. It provides judges, and others who interpret the law, with a model they might use in identifying the approach they have been using, and self-consciously to inspect and improve that model. A thoughtful judge might establish for himself, for example, a rough "threshold" of fit which any interpretation of data must meet in order to be "acceptable" on the dimension of fit, and then suppose that if more than one interpretation of some part of the law meets this thresh-old, the choice among these should be made, not through further and more precise comparisons between the two along that dimension, but by choosing the interpretation which is "substantively" better, that is, which better promotes the political ideals he thinks correct.

. . . Suppose a judge faces, for the first time, the possibility of overruling a narrow rule followed for some time in his jurisdiction. Suppose, for example, that the courts have consistently held, since the issue was first raised, that lawyers may not be sued for negligence. Our judge believes that this rule is wrong and unjust, and that it is inconsistent in principle with the general rule allowing actions in negligence against other professional people like doctors and accountants. Suppose he can nevertheless find some putative principle, in which others find [merit] though he does not, which would justify the distinction the law has rawn. Like the principle, for example, that because lawyers owe obligations to the courts or to abstract justice it would be unfair to impose on them the legal obligation of due care to their clients. He must ask whether the best interpretation of the past includes *that* principle in spite of the fact that he himself would reject it.

Neither answer to this question will seem wholly attractive to him. If he holds that the law does include this putative principle, then this argument would present the law, including the past decisions about suits against lawyers as coherent; but he would then expose what he would believe to be a flaw in the substantive law. He would be supposing that the law includes a principle he believes is wrong, and therefore has no place in a just and wise system. If he decides that the law does not include the putative principle, on the other hand, then he can properly regard this entire line of cases about actions against lawyers as mistakes, and ignore or overrule them; but he then exposes a flaw in the record of a different sort, namely that past judges have acted in an unprincipled way, and a demerit in his own decision, that it treats the lawyer who loses the present case differently from how judges have treated other lawyers in the past. He must ask which is, in the end, the greater of these flaws; which way of reading the record shows it,

in the last analysis, in the better and which in the worse light.

It would be absurd to suppose that all the lawyers and judges of any common law community share some set of convictions from which a single answer to that question could be deduced. Or even that many lawyers or judges would have ready at hand some convictions of their own which could supply an answer without further ado. But it is nevertheless possible for any judge to confront issues like these in a principled way, and this is what naturalism demands of him. He must accept that in deciding one way rather than another about the force of a line of precedents, for example, he is developing a working theory of legal interpretation in one rather than another direction, and this must seem to him the right direction as a matter of political principle, not simply an appealing direction for the moment because he likes the answer it recommends in the immediate case before him. Of course there is, in this counsel, much room for deception, including self-deception. But in most cases it will be possible for judges to recognize when they have submitted some issue to the discipline this description requires and also to recognize when some other judge has not.

Let me recapitulate. Interpretation is not a mechanical process. Nevertheless, judges can form working styles of interpretation, adequate for routine cases, and ready for refinement when cases are not routine. These working styles will include what I called formal features. They will set out, impressionistically, an account of fit, and may characterize a threshold of fit an interpretation must achieve in order to be eligible. But they will also contain a substantive part, formed from the judge's background political morality, or rather that part of his background morality which has become articulate in the course of his career. Sometimes this heuristic distinction between fit and substantive justice, as dimensions of a successful interpretation, will itself seem problematic, and a judge will be forced to elaborate that distinction by reflecting further on the full set of the substantive and procedural political rights of citizens a just legal system must respect and serve. In this way any truly hard case develops as well as engages a judge's style of adjudication.

NOTE

1. R. Dworkin, *Taking Rights Seriously* 105–130 (1977).

REVIEW AND DISCUSSION QUESTIONS

1. How does Dworkin define "naturalism"?
2. What is the "chain novel"?
3. Describe the two dimensions of interpretation used in legal adjudication.
4. Discuss the example of lawyers' being exempt from negligence lawsuits, explaining how it illustrates Dworkin's theory of law.
5. In what sense is Dworkin's theory a kind of natural law theory?
6. Explain whether Dworkin is more of a realist or a formalist.
7. It is sometimes said that following Dworkin's recommendation would mean that judges have more discretion than they should have or than they in fact do have. Do you agree? How could Dworkin respond?

The Dependence of Morality on Law

Tony Honoré

Discussions of morality and law invariably assume that we can determine what morality requires independently of what the law requires. In this essay, Tony Honoré, a fellow of All Souls College, Oxford, challenges that assumption. He argues that morality by itself is like an outline from which necessary details are missing and that without law to fill the gaps, morality is only an incomplete guide to conduct. What Honoré (following Aquinas) calls "determination" is necessary to give content to our moral obligations, and in a modern legal system with a rational morality only law can provide that determination.

MORALITY AND LAW

. . . Much writing about morality and law rests on an unstated premise. The premise is that we can determine what morality requires independently of the law of a community. What is right or wrong, good or bad, can be settled by rational argument without recourse to the formal institutions that make up a society's laws and legal system. If morality is universal, the same for everyone, that is still more clearly the case. Laws vary from one society to another, morality does not. Arguments of principle can settle what conduct is morally right or wrong, good or bad. This autonomous morality then provides a blueprint against which to judge laws and legal systems. Do they come up to scratch when set against the blueprint?

A whole series of questions then emerge. Are moral requirements, or some of them, necessarily part of a society's law? How far should they be made part of it? If a law violates these requirements, is it invalid? Are legal obligations "obligations" in the sense in which moral obligations clearly are? Or are they obligations in a different sense, and with a different force? Over the forty-five years since Hart reopened the debate about law and morality at Harvard in 1957, these and cognate issues have been keenly contested. On all of them the last word remains to be said.

But is the underlying premise correct? Can what morality requires be settled independently of law? Or is the picture of morality and law as model and copy in some ways misleading? Without reverting to the view which saw morality as a form of legislation (something like the Ten Commandments), I shall put the case for thinking that in any complex society a viable morality must have a legal component. It is incomplete without law and cannot be taken seriously unless filled out by law. Some obligations can be spelled out only by law. . . . So some laws at least relate to morality in two ways: they form part of it and at the same time are open to moral criticism.

. . . Can there be a morality without a legal component? Certainly in communities which do not have formal institutions, there will be a co-operative morality independent of law. In these (perhaps imaginary) simpler societies there are no laws, and morality is made determinate mainly by convention. But in complex societies it is doubtful whether morality could dispense with law. So an alternative view is worth considering. According to this the resources of morality are commonly overestimated. Morality on its own is incomplete and cannot provide a viable guide to what we are required to do in particular situations. . . . For moral thinking to guide the conduct of people living in communities and belonging to political entities, as we all do, it needs to be filled out by various means, including formal institutions. The main formal institutions are laws and legal systems.

To say this is not just to assert that law influences moral opinion, as moral opinion influences law, though their mutual influence is undeniable. Moral argument against sex or race discrimination leads to laws banning discrimination, and these laws in turn reinforce, or sometimes erode, the opinion that discrimination is wrong. Nor is the point being made that, given the shortcomings of human nature, laws are needed to give an incentive to conform to moral requirements. Laws reassure those who are disposed to conform that it is worth doing so and deter those who are disposed not to conform. That too is

undisputed, and represents one way in which law is indispensable to a viable morality. But it shows only that some parts of morality cannot be enforced without law, not that what is right or wrong depends on what the law is. . . .

All these connections are worth exploring. They point to ways in which laws influence moral opinion or suggest how morality might be developed or its viability guaranteed. This essay is however directed to . . . law's role in regard to gaps in morality. . . . My argument is that over a range of cases there can be no way of determining the right course of action without a legal component. Even a society of well-disposed angels, uniformly anxious to do right, needs a system of laws in order to know the right thing to do.

On this view law is part of the morality of any complex society. It follows that moral criticism directed against those laws that are part of its morality must take account of an awkward fact. If those laws are not morally compelling the society will to that extent lack a viable morality. The picture of morality as a blueprint and law as a structure put up according to or in disregard of it is then seen to be misleading. Morality is more like an outline from which details are missing. Laws, along with conventions, fill many of these in. As the legislator, judge, or official does this the details are incorporated in the overall plan, unless they diverge so far as to be incompatible with it. But we should be slow to reach the conclusion that the details do not fit the blueprint. If we do the building may have to be pulled down.

MORALITY AS AN INCOMPLETE GUIDE

. . . Morality, in the sense . . . of a co-operative morality which guides and assesses conduct from the point of view of its impact on others, can refer to a number of related things. There are the mores of a community, practices which consist in or are deducible from observed behaviour. These may not be uniform. Then there are the moral opinions of a community, what people in it regard as right or wrong, good or bad. These also may diverge and may not coincide with its mores. Finally, there are moral views which stand up to rational scrutiny. These in turn may not agree either with the dominant mores or moral opinions of any actual community. Take abortion. One can distinguish the practices—how many pregnant women have or refuse abortions—from whether those practices are regarded as right or wrong by the community in which they occur, or part of it. Both are distinct from the further question whether abortion really is right or wrong, or ometimes right and sometimes wrong; or, for those who think there is no such thing as objective right or wrong, whether it is rationally defensible or indefensible.

Though morality in all three forms is to some extent incomplete, it is the third type of morality, critical morality, that most clearly calls for a legal supplement. . . . Thinking about right and wrong . . . must include the aim of guiding behaviour in real life. Given this aim critical morality is incomplete as a guide to conduct as it affects others. The reason is that rational moral criticism has to be based on values that can be defended as worth pursuing. But these values are so general that they do not by themselves determine how people should behave in a given instance. We can seldom proceed by a process of deduction from the values to the required behaviour.

Early in the history of philosophy Aristotle made this simple point. In a well-known passage near the beginning of the *Nicomachean Ethics* he says that politics, a term which in his use includes morality, admits only a limited degree of precision. The subjects it studies, such as noble conduct and justice, are variable and uncertain. . . . At the outset of his course of lectures he invites his students to think about the problem which later, with Aquinas, is called that of determination (*determinatio*). This is the problem of making determinate a value such as noble conduct or justice. The process by which this occurs is best called "determination" rather than concretization, an ugly term, or implementation, a misleading one. Determination entails, for example, that some mechanism be found for settling what justice requires in a given situation. The same is true with the necessary adjustment for all the other values reflected in morality.

While the choice between different determinations (for example different methods of traffic control) may sometimes be morally neutral, it does not follow that the choice between having determinate

rules and not having them is morally indifferent. That there should be a determination is often necessary to the existence of a viable morality. If a busy town has no traffic regulations its moral standing, so far as avoiding injury to others is concerned, is not on a level with that of a busy city with regulations which might have been different. In one drivers and pedestrians know within limits how to behave. In the other people must decide for themselves how best to avoid threatened collisions as they move around. Mistakes are multiplied, and the duty to avoid harming others will not be carried out as well as it would otherwise have been, given the same degree of goodwill.

The need for determinants of morality is particularly clear as regards obligations owed by members of a community to their community. Taxation affords a good example. According to most people's moral outlook members of a community should make a contribution to the expense of meeting collective needs. A morality which denied this would hardly count as co-operative. In a monetary economy the contribution has to be mainly in money, and takes the form of paying taxes. So members of a community have in principle a moral obligation to pay taxes. But this obligation is incomplete or, if one prefers, inchoate, apart from law. It has no real content until the amount or rate of tax is fixed by an institutional decision, by law. What amounts to a reasonable contribution is not otherwise determinable, since what is required is a co-ordinated scheme which can be defended as fair not merely in the aggregate amount it raises but in its distribution. Taxpayers cannot settle it for themselves, as people can within limits settle for themselves, say, the proper way of showing respect for the feelings of others. Apart from law no one has a moral obligation to pay any particular amount of tax. An obligation to pay an indeterminate sum is not an effective obligation; it requires only a disposition, not an action. So, apart from law no one has an effective obligation to pay tax. If there were a society in which morality was taken as a sufficient guide to conduct apart from law it would therefore not be viable. It would grind to a halt and disintegrate for lack of resources. For this crucial moral and political obligation, vital to the life of a complex community, morality depends on law in the sense that to create an effective obligation it must have recourse to law.

It does not follow that the level and incidence of taxation are immune from moral criticism. The law is not conclusive of the obligation to pay, since the amount or distribution of the tax may be open to criticism. Whether taxes should be uniform or progressive, direct or indirect, as low as possible or as high as necessary depends on considerations of justice as well as on ease of assessment and efficiency in collection. But justice cannot by itself create an obligation to pay any particular amount of tax. So even if the tax is open to criticism, the criticism has to be viewed in the light of the fact that there is no way of fulfilling the obligation to help support the community apart from paying tax. . . .

But even if a determination is needed to make some moral obligations effective, must it take the form of law? Morality, some think, has the resources to make its own determinations. Or they can be made in other ways, without recourse to law. Four possible mechanisms for this will be discussed: subsumption, calculation, convention and command. The first can be briefly passed over. Sometimes a particular instance can be subsumed under a wider principle. If, for example, morality requires one to keep promises, then a person who makes a promise must do what he promised. The moral principle, combined with the facts, gives rise to an obligation. But in many situations there is no such simple connection with a moral principle, because the principle involved, for example that we should act with due concern for others, is not specific enough.

The second possible determinant, calculation, is an internal resource of morality. One may argue that, while some moral theories do not yield determinate obligations, others enable us to fix our obligations without recourse to law. If morality consists in performing duties and respecting the rights of others, an institutional determination is needed to fix these duties and rights. But if morality consists in maximizing certain consequences, or indeed in maximizing any values at all, people can calculate for themselves what conduct will maximize the appropriate values. In this way, for example, assuming agreement on the consequences or other values to be maximized, everyone can discover what amount of tax he should pay.

But the suggested calculation is impracticable, at least as regards obligations owed by members of a community to that community. To calculate one's proper tax liability, each person would need to know the total sum required and the contribution that everyone else ought on a similar calculation to make. The reason why individual tax liability cannot be calculated in isolation is that a co-operative morality involves some commitment to distributive justice. But it is not feasible to make these complex calculations, quite apart from the well known difficulties of comparing and quantifying values. . . .

In the upshot, calculation is not a viable determinant alternative to law. The third possible alternative is convention. The "*nomos*" in which, according to Aristotle, noble conduct or justice seems to consist can mean either convention or law. The difference is that conventions are not institutional but grow up informally by custom, usage, or practice, including trial and error, when people try to coordinate their activities. Convention and law are certainly closely related; and convention can often supply the necessary element of determination. Thus in a simple society there might be a convention that every household must supply an able-bodied man to serve in time of war. But in a more complex society it is clear that moral and political obligations cannot be fully determined by this means. Convention is not an effective means of settling problems that have not come up before, or not often enough to give rise to a settled practice. Nor does recourse to convention suffice when differences of opinion prevent the emergence of a settled practice, or when a minority flout the general practice. In that case a law is needed. Moreover conventions need interpreting, and for moral reasons, interpretation must be impartial. So, at least at a political level, an institutional mechanism is called for by which some recognized authority is given power to interpret the convention. However the authority may be described, this constitutes a recourse to law, and the decisions of the authority form precedents. So while convention can partly fix the content of some moral obligations, it cannot provide for their interpretation. Particularly in a complex society, it cannot settle the content of others at all. . . .

I have dealt with calculation and convention as determinants alternative to law. The fourth possibility is resort to commands. The moral and political importance of the distinction between laws and commands has perhaps not been fully grasped. The term is used here in the narrow sense of orders which are not open to legal or moral challenge. If they are open to challenge, like the orders of a policeman or an army officer according to contemporary western notions of law, they are subordinate aspects of law and not something distinct from it.

It would seem at first sight that unchallengeable commands might be the sole determinants of a morality which treated unquestioning obedience to superiors as a virtue of overriding importance. Perhaps no rational moral theory takes this form, but as some have thought otherwise, it is better not to cut short the discussion by too narrow a definition of co-operative morality. Given a moral theory, then, of this sort and a community with widely shared moral opinions, could their determinants take the form of commands issued by those in authority to give effect to prevailing moral opinion?

. . . A command-oriented moral theory of the sort just mentioned dominated the Third Reich. The theory was, in bare outline, composed of two elements along with a connecting link. The test of right and wrong was taken to be sound popular feeling, the moral opinions of the masses. These opinions were supposed to draw a sharp distinction between members and non-members (or unworthy members) of the community, for instance Aryans and non-Aryans. As regards members the morality was co-operative, as regards non-members hostile. But this morality in both its aspects was given effect by means of a second principle, the Leader Principle, which laid down that the commands of superiors be obeyed without question, whatever their content. It was not for the recipient of a command to judge whether it conformed to the test of sound popular feeling. These elements of Nazi morality were linked by ensuring that those in authority, entitled to issue binding commands, were committed to the morality of sound popular feeling.

This enabled Hitler, in principle at least, to substitute commands for laws as an instrument of government. Indeed he seldom resorted to legislation and was reluctant to embody in legislative

form the regime's most radical policies. For example the compulsory euthanasia programme, by which some seventy thousand people who were terminally ill or suffering from disabilities were put to death, was carried out on the basis of verbal orders. These ran counter to the Criminal Code, whose provisions as regards homicide remained in force throughout the period. But when this was pointed out Hitler rejected the suggestion that the programme should be put on a statutory basis. . . . More strikingly still, the holocaust itself, the Final Solution, lacked any legal basis; it was launched and carried through by verbal orders passed down from the top. It too was a violation on a grand scale of the Criminal Code.

This example suggests that for a rather special type of moral theory commands, in the narrow sense of orders not open to moral or legal challenge, can serve as a determinant of behaviour to the exclusion of law. But it also suggests that resort to command as a substitute for law is most likely to occur when those who have power are ashamed openly to declare how they propose to give effect to the morality they espouse.

. . . Commands can certainly be determinants of a morality of hostility. Indeed the substitution of command for law in the treatment of those regarded as enemies is a normal feature of oppressive regimes. How far commands can feature as determinants of a morality of co-operation is more debatable. Perhaps we should distinguish, within schemes of co-operation, between the rational and the non-rational. A morality which demands unquestioning obedience to commands is not a rational morality in the sense in which in the western tradition law is rational. It does not present those whom it addresses with genuine, though weighted, choices. It is a morality which works, and perhaps may succeed in ensuring co-operation, primarily (and not, as law does, secondarily) by instilling fear and foreclosing choice. If this is correct, then it is only in the context of a rational morality of co-operation that there is a sharp distinction between commands and laws. The determinants of such a rational morality must include laws. . . .

It seems, then, that neither calculation, convention nor command can wholly displace law as the determinant of a rational co-operative morality. But does this apply to the whole of morality, or only to certain parts of it? . . . In the upshot the minimum parts of morality that call for determination are those which make conduct compulsory. The legal determinants of these need not however necessarily impose obligations or duties. Many laws are in no way normative. They serve rather to define the conditions in which conduct is required or optional and to provide facilities for those who want to take advantage of the options open. For instance the duty to support one's family calls for determinants. But these need not relate directly to rights of support. They may include laws which confer on those who have a duty to support their families the power to make wills or set up trusts. Again the duty to protect others may call for legal determinants which include the power in certain conditions to punish wrongdoers. Moreover the need for determination varies from one sort of obligation to another. The duties of members of a community to that community can hardly be fixed except by law. That is less true of the obligations of members of a community, or of states in the international community, to one another. Here the difficulty is not, as with obligations owed to a community by its members, that no one's obligation can be determined unless everyone's is. On the contrary, abstract reasoning or convention will often indicate within limits what is required of each individual or state by way of, say, consideration for others. But complex problems of co-ordination cannot be solved in this way. For example respect for others and for efficiency requires that the use of wavelengths for transmitting information be co-ordinated. Interference both within and between states must be minimized. But no person or state has an obligation to refrain from using a particular wavelength unless wavelengths are allocated by municipal or international law. The same is true of other complex problems of co-ordination, for example in running transport, promoting health, or preserving the environment. In regard to these there is often no possibility of doing the right thing apart from law since there is, apart from law, no right thing to do.

It seems clear, then, that legal determinants of moral obligations are necessary to a co-operative morality whenever complex problems of co-ordination arise. Morality is not fully autonomous.

REVIEW AND DISCUSSION QUESTIONS

1. Would you agree that almost everyone takes for granted that we can determine the requirements of morality independent of law? If so, what explains this?
2. What are the three types of morality that Honoré distinguishes?
3. Using taxation as an example, explain what Honoré means by "determination" and why morality needs it. Are you persuaded that morality depends in an important way on law? Explain why or why not.
4. Why can't calculation, rather than law, provide the determination that morality needs? What about convention? The example of the Third Reich suggests that command might substitute for law in providing determination. How does Honoré deal with this example? Has Honoré overlooked any other possible non-legal sources of determination?
5. Are there important aspects of morality that don't require determination? In what ways is morality independent of law?
6. What are some of the other connections that exist between morality and law besides the dependence of morality on law that Honoré focuses on?

CHAPTER 10

International Law

From questions of war and peace to day-to-day commerce, in myriad ways international law structures and guides the interactions of nations. Yet the world has neither an international legislature nor an international court with compulsory jurisdiction. Because international law differs in these and other ways from national or municipal law, many theorists have been skeptical about its status as law. Is international law genuine law? In his contribution to this section, H. L. A. Hart addresses these doubts and affirms the reality of international law as law. Among the American legal cases involving questions of international law, its relation to morality, and the principles governing relations between sovereign states, none are more fascinating than the antebellum cases involving the Atlantic slave trade. A selection of these cases constitutes the second reading. Closer to our own time, the post–World War II Nuremberg trials of Nazi war criminals are widely seen as marking a genuine advance in international law because they held individuals accountable to a source of authority higher than their country's legal system or government. But the trials raise intriguing and important questions about the nature and status of these legal duties, about the responsibilities of individuals for their conduct, and about the limits of "victors' justice."

International Law

H. L. A. Hart

In this extract from his book *The Concept of Law,* H. L. A. Hart applies his legal philosophy (presented in Section 9 earlier) to international law, addressing two sources of doubt as to its status as genuine law. The first of these rests on the fact that international law lacks the organized sanctions that national or municipal law has for dealing with lawbreakers. The second doubt concerns whether it is possible for a state that is truly sovereign to be bound by or have genuine obligations under international law.

We shall consider two principal sources of doubt concerning the legal character of international law and, with them, the steps which theorists have taken to meet these doubts. Both forms of doubt arise from an adverse comparison of international law with municipal law, which is taken as the clear, standard example of what law is. The first has its roots deep in the conception of law as fundamentally a matter of orders backed by threats and contrasts the character of the *rules* of international law with those of municipal law. The second form of doubt springs from the obscure belief that states are fundamentally incapable of being the subjects of legal obligation, and contrasts the character of the *subjects* of international law with those of municipal law.

OBLIGATIONS AND SANCTIONS

The doubts which we shall consider are often expressed in the opening chapters of books on international law in the form of the question "How can international law be binding?" Yet there is something very confusing in this favourite form of question; and before we can deal with it we must face a prior question to which the answer is by no means clear. This prior question is: what is meant by saying of a whole system of law that it is "binding"? The statement that a particular rule of a system is binding on a particular person is one familiar to lawyers and tolerably clear in meaning. We may paraphrase it by the assertion that the rule in question is a valid rule, and under it the person in question has some obligation or duty. Besides this, there are some situations in which more general statements of this form are made. We may be doubtful in certain circumstances whether one legal system or another applies to a particular person. Such doubts may arise in the conflict of laws or in public international law. We may ask, in the former case, whether French or English Law is binding on a particular person as regards a particular transaction, and in the latter case we may ask whether the inhabitants of, for example, enemy-occupied Belgium, were bound by what the exiled government claimed was Belgian law or by the ordinances of the occupying power. But in both these cases, the questions are questions of law which arise *within* some system of law (municipal or international) and are settled by reference to the rules or principles of that system. They do not call in question the general character of the rules, but only their scope or applicability in given circumstances to particular persons or transactions. Plainly the question, "Is international law binding?" and its congeners "How can international law be binding?" or "What makes international law binding?" are questions of a different order. They express a doubt not about the applicability, but about the general legal status of international law: this doubt would be more candidly expressed in the form "Can such rules as these be meaningfully and truthfully said ever to give rise to obligations?" As the discussions in the books show, one source of doubt on this point is simply the absence from the system of centrally organized sanctions. This is one point of adverse comparison with municipal law, the rules of which are taken to be unquestionably "binding" and to be paradigms of legal obligation. From this stage the further argument is simple: if for this reason the rules of international law are not "binding", it is surely indefensible to take seriously their classification as law; for however tolerant the modes of common speech may be, this is too great a difference to be overlooked. All speculation about the nature of law begins from the assumption that its existence at least makes certain conduct obligatory. . . .

To argue that international law is not binding because of its lack of organized sanctions is tacitly to accept the analysis of obligation contained in the theory that law is essentially a matter of orders backed by threats. This theory, as we have seen, identifies "having an obligation" or "being bound" with "likely to suffer the sanction or punishment threatened for disobedience" Yet, as we have argued, this identification distorts the role played in all legal thought and discourse of the ideas of obligation and duty. Even in municipal law, where there are effective organized sanctions, we must distinguish . . . the meaning of the external predictive statement "I (you) are likely to suffer for disobedience" from the internal normative statement "I (you) have an obligation to act thus" which assesses a particular person's situation from the point of view of rules accepted as guiding standards of behaviour. It is true that not all rules give rise to

obligations or duties; and it is also true that the rules which do so generally call for some sacrifice of private interests, and are generally supported by serious demands for conformity and insistent criticism of deviations. Yet once we free ourselves from the predictive analysis and its parent conception of law as essentially an order backed by threats, there seems no good reason for limiting the normative idea of obligation to rules supported by organized sanctions.

We must, however, consider another form of the argument, more plausible because it is not committed to definition of obligation in terms of the likelihood of threatened sanctions. The sceptic may point out that there are in a municipal system, as we have ourselves stressed, certain provisions which are justifiably called necessary; among these are primary rules of obligation, prohibiting the free use of violence, and rules providing for the official use of force as a sanction for these and other rules. If such rules and organized sanctions supporting them are in this sense necessary for municipal law, are they not equally so for international law? . . .

The answer to the argument in this form is to be found in those elementary truths about human beings and their environment which constitute the enduring psychological and physical setting of municipal law. In societies of individuals, approximately equal in physical strength and vulnerability, physical sanctions are both necessary and possible. They are required in order that those who would voluntarily submit to the restraints of law shall not be mere victims of malefactors who would, in the absence of such sanctions, reap the advantages of respect for law on the part of others, without respecting it themselves. . . . Yet, because of the same fact of approximate equality and the patent advantages of submission to a system of restraints, no combination of malefactors is likely to exceed in strength those who would voluntarily co-operate in its maintenance. In these circumstances, which constitute the background of municipal law, sanctions may successfully be used against malefactors with relatively small risks, and the threat of them will add much to whatever natural deterrents there may be. But, just because the simple truisms which hold good for individuals do not hold good for states, and the factual background to international law is so different from that of municipal law, there is neither a similar necessity for sanctions (desirable though it may be that international law should be supported by them) nor a similar prospect of their safe and efficacious use.

This is so because aggression between states is very unlike that between individuals. . . . To initiate a war is, even for the strongest power, to risk much for an outcome which is rarely predictable with reasonable confidence. On the other hand, because of the inequality of states, there can be no standing assurance that the combined strength of those on the side of international order is likely to preponderate over the powers tempted to aggression. Hence the organization and use of sanctions may involve fearful risks and the threat of them add little to the natural deterrents. Against this very different background of fact, international law has developed in a form different from that of municipal law. In a population of a modern state, if there were no organized repression and punishment of crime, violence and theft would be hourly expected; but for states, long years of peace have intervened between disastrous wars. These years of peace are only rationally to be expected, given the risks and stakes of war and the mutual needs of states; but they are worth regulating by rules which differ from those of municipal law in (among other things) not providing for their enforcement by any central organ. Yet what these rules require is thought and spoken of as obligatory; there is general pressure for conformity to the rules; claims and admissions are based on them and their breach is held to justify not only insistent demands for compensation, but reprisals and countermeasures. When the rules are disregarded, it is not on the footing that they are not binding; instead efforts are made to conceal the facts. It may of course be said that such rules are efficacious only so far as they concern issues over which states are unwilling to fight. This may be so, and may reflect adversely on the importance of the system and its value to humanity. Yet that even so much may be secured shows that no simple deduction can be made from the necessity of organized sanctions to municipal law, in its setting of physical and psychological facts, to the conclusion that without them international law, in its very different setting, imposes no obligations, is not "binding," and so not worth the title of "law."

OBLIGATION AND THE SOVEREIGNTY OF STATES

. . . One of the most persistent sources of perplexity about the obligatory character of international law has been the difficulty felt in accepting or explaining the fact that a state which is sovereign may also be "bound" by, or have an obligation under, international law. This form of scepticism is, in a sense, more extreme than the objection that international law is not binding because it lacks sanctions. For whereas that would be met if one day international law were reinforced by a system of sanctions, the present objection is based on a radical inconsistency, said or felt to exist, in the conception of a state which is at once sovereign and subject to law.

Examination of this objection involves a scrutiny of the notion of sovereignty, applied not to a legislature or to some other element or person *within* a state, tub to a state itself. Whenever the word "sovereign" appears in jurisprudence, there is a tendency to associate with it the idea of a person above the law whose word is law for his inferiors or subjects. . . .

States such as Great Britain or Brazil, the United States or Italy . . . possess a very large measure of independence from both legal and factual control by any authorities or persons outside their borders, and would rank as "sovereign states" in international law. On the other hand, individual states which are members of a federal union, such as the United States, are subject in many different ways to the authority and control of the federal government and constitution. Yet the independence which even these federated states retain is large if we compare it with the position, say, of an English county, of which the word "state" would not be used at all. A county may have a local council discharging, for its area, some of the functions of a legislature, but its meagre powers are subordinate to those of Parliament and, except in certain minor respects, the area of the county is subject to the same laws and government as the rest of the country.

Between these extremes there are many different types and degrees of dependence (and so of independence) between territorial units which possess an ordered government. Colonies, protectorates, suzerainties, trust territories, confederations, present fascinating problems of classification from this point of view. . . . It is possible to imagine many different forms of international authority and correspondingly many different limitations on the independence of states. The possibilities include, among many others, a world legislature on the model of the British Parliament, possessing legally unlimited powers to regulate the internal and external affairs of all; a federal legislature on the model of Congress, with legal competence only over specified matters or one limited by guarantees of specific rights of the constituent units; a régime in which the only form of legal control consists of rules generally accepted as applicable to all; and finally a régime in which the only form of obligation recognized is contractual or self-imposed, so that a state's independence is legally limited only by its own act.

It is salutary to consider this range of possibilities because merely to realize that there are many possible forms and degrees of dependence and independence, is a step towards answering the claim that because states are sovereign they "*cannot*" be subject to or bound by international law or "*can*" only be bound by some specific form of international law. For the word "sovereign" means here no more than "independent"; and, like the latter, is negative in force: a sovereign state is one *not* subject to certain types of control, and its sovereignty is that area of conduct in which it is autonomous. . . . Hence we can only know which states are sovereign, and what the extent of their sovereignty is, when we know what the rules are; just as we can only know whether an Englishman or an American is free and the extent of his freedom when we know what English or American law is. The rules of international law are indeed vague and conflicting on many points, so that doubt about the area of independence left to states is far greater than that concerning the extent of a citizen's freedom under municipal law.

. . . The question for municipal law is: what is the extent of the supreme legislative authority recognized in this system? For international law it is: what is the maximum area of autonomy which the rules allow to states?

Thus the simplest answer to the present objection is that it inverts the order in which questions must be considered. There is no way of knowing what sovereignty states have, till we know what the forms of international law are and whether or not

they are mere empty forms. Much juristic debate has been confused because this principle has been ignored, and it is profitable to consider in its light those theories of international law which are known as "voluntarist" or theories of "auto-limitation." These attempted to reconcile the (absolute) sovereignty of states with the existence of binding rules of international law, by treating all international obligations as self-imposed like the obligation which arises from a promise. . . .

The . . . view that a state may impose obligations on itself by promise, agreement, or treaty is not, however, consistent with the theory that states are subject only to rules which they have thus imposed on themselves. For, in order that words, spoken or written, should in certain circumstances function as a promise, agreement, or treaty, and so give rise to obligations and confer rights which others may claim, *rules* must already exist providing that a state is bound to do whatever it undertakes by appropriate words to do. Such rules presupposed in the very notion of a self-imposed obligation obviously cannot derive *their* obligatory status from a self-imposed obligation to obey them. . . . What is necessary and sufficient, in order that the words of a promise, agreement, or treaty should give rise to obligations, is that rules providing for this and specifying a procedure for these self-binding operations should be generally . . . acknowledged. Where they are acknowledged the individual or state who wittingly uses these procedures is bound thereby, whether he or it chooses to be bound or not. Hence, even this most voluntary form of social obligation involves some rules which are binding independently of the choice of the party bound by them, and this, in the case of states, is inconsistent with the supposition that their sovereignty demands freedom from all such rules. . . .

A detailed scrutiny of the claim that all international obligation arises from the consent of the party bound cannot be undertaken here, but two clear and important exceptions to this doctrine must be noticed. The first is the case of a new state. It has never been doubted that when a new, independent state emerges into existence, as did Iraq in 1932, and Israel in 1948, it is bound by the general obligations of international law including, among others, the rules that give binding force to treaties. Here the attempt to rest the new state's international obligations on a "tacit" or "inferred" consent seems wholly threadbare. The second case is that of a state acquiring territory or undergoing some other change, which brings with it, for the first time, the incidence of obligations under rules which previously it had no opportunity either to observe or break, and to which it had no occasion to give or withhold consent. If a state, previously without access to the sea, acquires maritime territory, it is clear that this is enough to make it subject to all the rules of international law relating to the territorial waters and the high seas. . . .

It has often been said in the discussion of the "foundations" of international law, that in the last resort, the rules of international law *must* rest on the conviction of states that there is a moral obligation to obey them; yet, if this means more than that the obligations which they recognize are not enforceable by officially organized sanctions, there seems no reason to accept it. Of course it is possible to think of circumstances which would certainly justify our saying that a state considered some course of conduct required by international law morally obligatory, and acted for that reason. It might, for example, continue to perform the obligations of an onerous treaty because of the manifest harm to humanity that would follow if confidence in treaties was severely shaken, or because of the sense that it was only fair to shoulder the irksome burdens of a code from which it, in its turn, had profited in the past when the burden fell on others. Precisely whose motives, thoughts and feelings on such matters of moral conviction are to be attributed to the state is a question which need not detain us here.

But though there *may* be such a sense of moral obligation it is difficult to see why or in what sense it *must* exist as a condition of the existence of international law. It is clear that in the practice of states certain rules are regularly respected even at the cost of certain sacrifices; claims are formulated by reference to them; breaches of the rules expose the offender to serious criticism and are held to justify claims for compensation or retaliation. These, surely, are all the elements required to support the statement that there exist among states rules imposing obligations upon them. The proof that "binding" rules in any society exist, is simply that they are thought of, spoken of, and function as such. What more is required by way of "foundations" and why,

if more is required, must it be a foundation of moral obligation? It is, of course, true that rules could not exist or function in the relations between states unless a preponderant majority accepted the rules and voluntarily co-operated in maintaining them. It is true also that the pressure exercised on those who break or threaten to break the rules is often relatively weak, and has usually been decentralized or unorganized. But as in the case of individuals, who voluntarily accept the far more strongly coercive system of municipal law, the motives for voluntarily supporting such a system may be extremely diverse. It may well be that any form of legal order is at its healthiest when there is a generally diffused sense that it is morally obligatory to conform to it. Nonetheless, adherence to law may not be motivated by it, but by calculations of long-term interest, or by the wish to continue a tradition or by disinterested concern for others. There seems no good reason for identifying any of these as a necessary condition of the existence of law either among individuals or states.

REVIEW AND DISCUSSION QUESTIONS

1. How does Hart respond to the argument that international law is not binding because it lacks organized sanctions?
2. How does "the factual background to international law" (in Hart's words) differ from that of municipal law? Is Hart right to maintain that these differences fail to establish either that international law is not binding or that it is not really law?
3. Is there a contradiction between being a sovereign state and being bound by international law? Explain why or why not. Hart says that we can't know "what sovereignty states have till we know what the forms of international law are." Explain.
4. Explain how Hart responds to the argument that the rules of international law can only be binding on states if they are self-imposed (like the obligations that arise from making a promise).
5. Hart disputes the idea that "the rules of international law *must* rest on the conviction of states that they have a moral obligation to obey them." Explain and critically assess his reasoning.
6. Can international law be squared with Hart's model of law as the union of primary and secondary rule (see Section 9)? In your view is international law a genuine form of law? Explain.

The African Slave Trade

United States v. La Jeune Eugenie; The Antelope; and *The Amistad*

Before the Civil War, the U.S. Constitution permitted slavery, the legality of which was left for each individual state to determine for itself. In 1808, however, Congress made it illegal to import slaves into the country, and it subsequently passed laws punishing people for engaging in the African slave trade. As a result, on several occasions, federal courts had to decide on the fate of foreign slave-trading vessels that had fallen into the hands of U.S. officials and of any slaves found onboard them. All of these cases touch not only on the morality of slavery, but also on the nature and requirements of international law and the mutual obligations among states.

In the *United States v. Le Jeune Eugenie* (1822), Supreme Court Justice Story, presiding over a case before the U.S. Circuit Court, ruled that the government had the authority to confiscate a French slave-trading vessel it had captured because the Atlantic slave trade violated natural law. However, in the

26 F. Cas. 832, 846 (C.C.C. Mass. 1822, No. 15,551); 23 U.S. (10 Wheat.), 66, 120–22 (1825); and 40 U.S. 518 (1841).

Antelope case (1825), in an opinion written by Chief Justice Marshall, the Supreme Court ruled against the United States, which had argued that Africans seized by American privateers from Spanish and Portuguese vessels and brought to the U.S. should be set free. The famous *Amistad* case, brought before the Supreme Court in 1841, concerned the fate of a group of Africans who had been abducted and then sold to two Spaniards. The Africans subsequently seized the ship they were on, killed two of the crew, and ordered the remaining crew to sail to Africa. The crew, however, sailed to New York, where the Africans were imprisoned.

United States v. La Jeune Eugenie

Justice Story delivered the opinion of the court: Now the law of nations may be deduced, first, from the general principles of right and justice, applied to the concerns of individuals, and thence to the relations and duties of nations; or, secondly, in things indifferent or questionable, from the customary observances and recognitions of civilized nations; or, lastly, from the conventional or positive law, that regulates the intercourse between states. . . . But I think it may be unequivocally affirmed, that every doctrine, that may be fairly deduced by correct reasoning from the rights and duties of nations, and the nature of moral obligation, may theoretically be said to exist in the law of nations; and unless it be relaxed or waived by the consent of nations, which may be evidenced by their general practice and customs, it may be enforced by a court of justice, whenever it arises in judgment. And I may go further and say, that no practice whatsoever can obliterate the fundamental distinction between right and wrong, and that every nation is at liberty to apply to another the correct principle, whenever both nations by their public acts recede from such practice and admit the injustice or cruelty of it.

Now in respect to the African slave trade . . . it cannot admit of serious question, that it is founded in a violation of some of the first principles, which ought to govern nations. It is repugnant to the great principles of Christian duty, the dictates of natural religion, the obligations of good faith and morality, and the eternal maxims of social justice. When any trade can be truly said to have these ingredients, it is impossible, that it can be consistent with any system of law, that purports to rest on the authority of reason or revelation. And it is sufficient to stamp any trade as interdicted by public law, when it can be justly affirmed, that it is repugnant to the general principles of justice and humanity.

The Antelope

Marshall, Chief Justice: In prosecuting this appeal, the United States assert no property in themselves. They appear in the character of guardians, or next friends, of these Africans, who are brought, without any act of their own, into the bosom of our country, insist on their right to freedom, and submit their claim to the laws of the land, and to the tribunals of the nation. The consuls of Spain and Portugal, respectively, demand these Africans as slaves, who have, in the regular cause of legitimate commerce, been acquired as property, by the subjects of their respective sovereigns, and claim their restitution under the laws of the United States. . . .

That the course of opinion on the slave-trade should be unsettled, ought to excite no surprise. The Christian and civilized nations of the world with whom we have most intercourse, have all been engaged in it. However abhorrent this traffic may be to a mind whose original feelings are not blunted by familiarity with the practice, it has been sanctioned, in modern times, by the laws of all nations who possess distant colonies, each of whom has engaged in it as a common commercial

business, which no other could rightfully interrupt. It has claimed all the sanction which could be derived from long usage and general acquiescence. That trade could not be considered as contrary to the law of nations which was authorized and protected by the laws of all commercial nations; the right to carry on which was claimed by each, and allowed by each.

. . . In the beginning of this century, several humane and enlightened individuals of Great Britain devoted themselves to the cause of the Africans; and by frequent appeals to the nation, in which the enormity of this commerce was unveiled and exposed to the public eye, the general sentiment was at length roused against it, and the feelings of justice and humanity, regaining their long-lost ascendency, prevailed so far in the British parliament, as to obtain an act for its abolition. The utmost efforts of the British government, as well as of that of the United States, have since been assiduously employed in its suppression. It has been denounced by both, in terms of great severity, and those concerned in it are subjected to the heaviest penalties which law can inflict. In addition to these measures, operating on their own people, they have used all their influence to bring other nations into the same system, and to interdict this trade by the consent of all. . . . While its illegality is asserted by some governments, but not admitted by all . . . it is not wonderful . . . that opposite opinions should be entertained on the precise cases in which our own laws may control and limit the practice of others. . . .

The question, whether the slave trade is prohibited by the law of nations has been seriously propounded, and both the affirmative and negative of the proposition have been maintained with equal earnestness.

That it is contrary to the law of nature will scarcely be denied. That every man has a natural right to the fruits of his own labour, is generally admitted; and that no other person can rightfully deprive him of those fruits, and appropriate them against his will, seems to be the necessary result of this admission. But from the earliest times war has existed, and war confers rights in which all have acquiesced. Among the most enlightened nations of antiquity, one of these was, that the victor might enslave the vanquished. This, which was the usage of all, could not be pronounced repugnant to the law of nations, which is certainly to be tried by the test of general usage. That which has received the assent of all, must be the law of all.

Slavery, then, has its origin in force; but as the world has agreed that it is a legitimate result of force, the state of things which is thus produced by general consent, cannot be pronounced unlawful.

Throughout Christendom, this harsh rule has been exploded, and war is no longer considered as giving a right to enslave captives. But this triumph of humanity has not been universal. The parties to the modern law of nations do not propagate their principles by force; and Africa has not yet adopted them. Throughout the whole extent of that immense continent, so far as we know its history, it is still the law of nations that prisoners are slaves. Can those who have themselves renounced this law, be permitted to participate in its effects by purchasing the beings who are its victims?

Whatever might be the answer of a moralist to this question, a jurist must search for its legal solution, in those principles of action which are sanctioned by the usages, the national acts, and the general assent, of that portion of the world of which he considers himself as a part, and to whose law the appeal is made. If we resort to this standard as the test of international law, the question, as has already been observed, is decided in favour of the legality of the trade. Both Europe and America embarked on it; and for nearly two centuries, it was carried on without opposition, and without censure. A jurist could not say, that a practice thus supported was illegal, and that those engaged in it might be punished, either personally, or by deprivation of property.

In this commerce, thus sanctioned by universal assent, every nation had an equal right to engage. How is this right to be lost? Each may renounce it for its own people; but can this renunciation affect others?

No principle of general law is more universally acknowledged, than the perfect equality of nations. Russia and Geneva have equal rights. It results from this equality, that no one can rightfully impose a rule on another. Each legislates for itself, but its legislation can operate on itself alone. A right, then, which is vested in all by the consent of all, can be divested only by consent; and this

trade, in which all have participated, must remain lawful to those who cannot be induced to relinquish it. As no nation can prescribe a rule for others, none can make a law of nations; and this traffic remains lawful to those whose governments have not forbidden it.

The Amistad

Mr. Justice Story delivered the opinion of the Court: Before entering upon the discussion of the main points involved in this interesting and important controversy, it may be necessary to say a few words as to the actual posture of the case as it now stands before us. In the first place, then, the only parties now before the Court on one side are the United States, intervening for the sole purpose of procuring restitution of the property as Spanish property, pursuant to a treaty with Spain.... The United States do not assert any property in themselves, or any violation of their own rights, or sovereignty, or laws, by the acts complained of. They do not insist that these negroes have been imported into the United States, in contravention of our own slave trade acts. They do not seek to have these negroes delivered up for the purpose of being transported to Cuba as pirates or robbers, or as fugitive criminals against the laws of Spain.... They simply confine themselves to the right of the Spanish claimants to the restitution of their property....

In the next place, the parties before the Court on the other side as appellees are ... the negroes (Cinque, and others), asserting themselves, in their answer, not to be slaves, but free native Africans, kidnapped in their own country, and illegally transported by force from that country; and now entitled to maintain their freedom.

No question has been here made, as to the proprietary interests in the vessel and cargo. It is admitted that they belong to Spanish subjects, and that they ought to be restored.... The main controversy is, whether these negroes are the property of Ruiz and Montez, and ought to be delivered up; and to this, accordingly, we shall first direct our attention.

It has been argued on behalf of the United States, that the Court is bound to deliver them up, according to the treaty of 1795, with Spain, which has in this particular been continued in full force, by the treaty of 1819, ratified in 1821.... The ninth article provides, "that all ships and merchandise, of what nature soever, which shall be rescued out of the hands of any pirates or robbers, on the high seas, shall be brought into some port of either state, and shall be delivered to the custody of the officers of that port, in order to be taken care of and restored entire to the true proprietor, as soon as due and sufficient proof shall be made concerning the property thereof." This is the article on which the main reliance is placed on behalf of the United States, for the restitution of these negroes. To bring the case within the article, it is essential to establish, First, That these negroes, under all the circumstances, fall within the description of merchandise, in the sense of the treaty. Secondly, That there has been a rescue of them on the high seas, out of the hands of the pirates and robbers; which, in the present case, can only be, by showing that they themselves are pirates and robbers; and, Thirdly, That Ruiz and Montez, the asserted proprietors, are the true proprietors, and have established their title by competent proof.

If these negroes were, at the time, lawfully held as slaves under the laws of Spain, and recognised by those laws as property capable of being lawfully bought and sold, we see no reason why they may not justly be deemed within the intent of the treaty, to be included under the denomination of merchandise, and, as such, ought to be restored to the claimants.... But, admitting this, it is clear, in our opinion, that neither of the other essential facts and requisites has been established in proof....

It is plain beyond controversy, if we examine the evidence, that these negroes never were the lawful slaves of Ruiz or Montez, or of any other Spanish subjects. They are natives of Africa, and were kidnapped there, and were unlawfully transported to Cuba, in violation of the laws and treaties of Spain.... By those laws, and treaties, and edicts, the African slave trade is utterly abolished; the dealing in that trade is deemed a heinous crime; and the negroes thereby introduced into the dominions of Spain, are declared to be free....

If, then, these negroes are not slaves, but are kidnapped Africans, who, by the laws of Spain itself,

are entitled to their freedom, and were kidnapped and illegally carried to Cuba, and illegally detained and restrained on board of the Amistad, there is no pretence to say, that they are pirates or robbers. We may lament the dreadful acts, by which they asserted their liberty, and took possession of the Amistad, and endeavoured to regain their native country; but they cannot be deemed pirates or robbers in the sense of the law of nations, or the treaty with Spain, or the laws of Spain itself. . . .

It is also a most important consideration in the present case, which ought not to be lost sight of, that, supposing these African negroes not to be slaves, but kidnapped, and free negroes, the treaty with Spain cannot be obligatory upon them; and the United States are bound to respect their rights as much as those of Spanish subjects. The conflict of rights between the parties under such circumstances, becomes positive and inevitable, and must be decided upon the eternal principles of justice and international law. If the contest were about any goods on board of this ship, to which American citizens asserted a title, which was denied by the Spanish claimants, there could be no doubt of the right of such American citizens to litigate their claims before any competent American tribunal, notwithstanding the treaty with Spain. A fortiori, the doctrine must apply where human life and human liberty are in issue; and constitute the very essence of the controversy. The treaty with Spain never could have intended to take away the equal rights of all foreigners, who should contest their claims before any of our Courts, to equal justice; or to deprive such foreigners of the protection given them by other treaties, or by the general law of nations. Upon the merits of the case, then, there does not seem to us to be any ground for doubt, that these negroes ought to be deemed free; and that the Spanish treaty interposes no obstacle to the just assertion of their rights.

There is another consideration growing out of this part of the case, which necessarily rises in judgment. It is observable, that the United States, in their original claim, filed it in the alternative, to have the negroes, if slaves and Spanish property, restored to the proprietors; or, if not slaves, but negroes who had been transported from Africa, in violation of the laws of the United States, and brought into the United States contrary to the same laws, then the Court to pass an order to enable the United States to remove such persons to the coast of Africa, to be delivered there to such agent as may be authorized to receive and provide for them. . . . The decree of the District Court . . . contained an order for the delivery of the negroes to the United States, to be transported to the coast of Africa, under the act of the 3d of March, 1819, ch. 224. . . . In our judgment, upon the admitted facts, there is no ground to assert that the case comes within the purview of the act of 1819, or of any other of our prohibitory slave trade acts. These negroes were never taken from Africa, or brought to the United States in contravention of those acts. When the Amistad arrived she was in possession of the negroes, asserting their freedom; and in no sense could they possibly intend to import themselves here, as slaves, or for sale as slaves. In this view of the matter, that part of the decree of the District Court is unmaintainable, and must be reversed. . . .

Upon the whole, our opinion is, that the decree of the Circuit Court, affirming that of the District Court, ought to be affirmed, except so far as it directs the negroes to be delivered to the President, to be transported to Africa, in pursuance of the act of the 3d of March, 1819; and, as to this, it ought to be reversed: and that the said negroes be declared to be free, and be dismissed from the custody of the Court, and go without delay.

REVIEW AND DISCUSSION QUESTIONS

1. On what principles does Justice Story base his decision in *La Jeune Eugenie?*
2. According to Story, "every nation is at liberty to apply to another [nation] the correct principle [of right and wrong]." What are the implications of this for international law?
3. In the *Antelope* case, Marshall concedes that the slave trade is contrary to the law of nature. On what grounds, then, does he justify ordering the slaves to be returned to the foreign slave traders? Why, in his view, does the slave trade not violate international law? Are you persuaded by his reasoning?

4. What does Marshall see as the test of international law? Contrast his implicit legal philosophy with that of Story in the previous case.
5. In the *Amistad* case, Justice Story states that, according to the treaty with Spain, three things must be established if the Africans are to be returned to Spain. What are they, and what does the Court say about each point?
6. Contrast Story's decision in *Amistad* with his earlier opinion in *La Jeune Eugenie.* In what way do you think his *Amistad* opinion was influenced by the precedent of *Antelope?* Do considerations of natural law still play a role in his legal reasoning?

International Law and Individual Responsibility

The International Military Tribunal at Nuremberg

After World War II, the victorious powers—the United States, Britain, France, and the Soviet Union—established an International Military Tribunal to try Germans accused of war crimes. Among those tried at Nuremberg were Germany's military leaders, the so-called German General Staff. The reading below contains three parts: the defense's argument on behalf of the military leaders, two statements for the prosecution, and the Tribunal's judgment, which explains the Charter that governs it and defends its legitimacy. The judgment lists the crimes for which the defendants are charged and upholds the principle that individuals have international duties that transcend national obligations of obedience.

DEFENSE ARGUMENTS

Dr. Hans Laternser: . . . Hitler, who was the head of the German Reich, and the Supreme Commander of the Armed Forces, has eluded judicial responsibility by his death. Since he can no longer be dealt with, the Prosecution have taken the highest military commanders instead of the Supreme Commander and head of the State, made them summarily also political leaders, and are attempting in this way to render them responsible.

This method is indeed unique and without precedent in the history of nations, and may well be contemplated with peculiar feelings by all soldiers of the world. . . .

If the Prosecutor, Colonel Taylor, argues that Hitler could not have waged his wars without the assistance of the Armed Forces, that argument cannot be invalidated. Nobody has ever been able to wage a war without soldiers. However, what Carlyle says is true for the German military leaders as for all soldiers:

> If a man becomes a soldier, his soul and his body thereby become the property of his commanding officer. He is not allowed to decide for himself whether the cause for which he fights is good or bad. His enemies are selected for him, and not by him. It is his duty to obey and to ask no questions.

. . . By indicting the military leader—who, obeying the orders of his government, has fulfilled his military duties—because the Prosecution declares the action of his government to be illegal and represents him as a partner to such action of the government, the Prosecution places upon him the obligation to examine the legality of his country's policy, and raises him to the position of a judge called upon to give a verdict on the policy of his state.

It cannot be my task to present the consequences of such a mental revolution for the soldiers of the world. . . .

Since Hitler is dead, the Prosecution leaves him in the background, and looks for other responsible parties. Yet no one can deny that Hitler alone wielded the power of the Reich in his hands, and consequently also had the sole and total responsibility. The essence of every dictatorship ultimately lies in the fact that one man's will is almighty, that his will is decisive in all matters. In no other dictatorship was this principle developed so exclusively as in Hitler's dictatorship. . . .

This picture should not be interpreted as an attempt to elude responsibility: every German general is enough of a man to stand up for his actions; but if justice is to be done, the actual circumstances, as they really were, must be recognized and serve as a basis for the final judgment. The best proof, however, against the participation of the generals in Hitler's plans is given by Hitler himself when he says: "I do not expect my generals to understand my orders; I only expect them to obey them." . . .

The last count of the Indictment, that the military leaders had rendered themselves guilty because they tolerated in practice Hitler's criminal plans and deeds, instead of revolting against them, returns us again to the central problem of these proceedings against the soldiers: the problem of the duty to obey. It has been repeatedly stated that the Führer order was not only a military order, but that it had, over and above this, a legislative effect.

Thus were not the military leaders simply bound to obey the law? If the duty to obey does not exist in the case of an order which aims at a civil crime, it is because the order demands an action directed against the authority of the State. But can there be any question of a crime if the order requires action which is not directed against the authority of the State, but on the contrary is demanded by that authority? And even if we reply to this question in the affirmative, what citizen of any country in the world is in a position to recognize the criminal nature of his action?

It is not sufficient, in order to ascertain guilt, that the Prosecution explain what the defendants should not have done—at the same time, they should tell us what they might, ought, and should have done, for any legal prohibition must also include a positive directive. If I suppose that, in spite of the sovereignty of the individual states, a legal obligation existed for the generals to act in accordance with international law and moral requirements even against the law of their own State, such a legal obligation could only be affirmed if the corresponding action offered a chance of success. After all, to allow oneself to be hanged merely to evade one's duties, to betray one's country without any prospect of being able to change matters—these things cannot be demanded by virtue of any morality. After all, there is no obligation for anybody to become a martyr.

And what were the possibilities of negative or positive action against orders and law on the part of the indicted generals? What were the chances of success? The simple rejection of unlawful plans or orders, be it by contradiction, warning, representations, objections, or the like, would have been in theory possible, but utterly unsuccessful in practice. . . .

A democratic politician might say that they could have resigned. That is a practical possibility for a parliamentary minister in a democratic country—a German officer could not resign. He was bound by his military oath, which was a supreme obligation for the veteran officer, more than for anyone else. A German general could only ask for approval of his resignation. Whether that request was successful or not was beyond his influence. Moreover, during the war, Hitler prohibited any such request, and placed resignation on the same footing as desertion. . . .

Let me recapitulate:

1. Military disobedience is and remains a violation of duty, in times of war a crime punishable with the death penalty.
2. There exists no duty to disobey for any soldier in the world, as long as states with a sovereignty of their own continue to exist.
3. Under Hitler's dictatorship, open disobedience would only have led to the destruction of the subordinate, but never to a repeal of orders given.
4. No class has made, through its highest representatives, such great sacrifices for its conceptions as opposed to Hitler's methods, as the circle of officers who are indicted before this Tribunal.

In view of the impossibility and the ineffectiveness of any passive resistance, there would have remained only the method of violence, rebellion, and *coup d'état.* Whoever contemplated this method had to be aware of the fact that it would have to involve the removal of Hitler and of the leading men of the Party in such a way that these men would be put to death. There was, therefore, at the beginning of each *coup d'état* the inexorable compulsion to liquidate Hitler and the leading men of the Party.

To the soldier this meant murder and disloyalty to his oath. Even if it is demanded that the generals, for reasons of a higher world morality, ought to have sacrificed their personal and military honor, how could they have been justified in taking such action

against the will of the Nation, and when could such action have been effected with good chances of success and for the benefit of the people? . . .

As citizens of the State they might have claimed for themselves the right to refuse service to a Führer and a system which, the longer the war lasted, proved to be more and more harmful. They might thus have evaded their personal responsibility. . . . Perhaps they would not now be before this Tribunal. But by taking such a decision they would at the same time have let down their soldiers who trusted them and for whom they felt responsible. Therefore there remained for them as soldiers only the duty to fight. This "duty" might, in a wider sense of the word, have consisted in overthrowing the system. In war, however, this would practically have amounted to nothing less than inviting defeat. No soldier could take that upon himself. No military leader can for years demand of his soldiers that they should give their lives, and then abandon his post himself and go down in history as a traitor to his nation. . . .

Thus there remained for the German military leaders only the duty to fight the enemy to the last. Confronted with the tragic decision between personal rights and soldierly duties, they decided in favor of their duties, and thus acted in the spirit of soldierly morale.

PROSECUTION ARGUMENTS

Sir David Maxwell-Fyfe: . . . The men involved have joined in wars which they knew were unjust wars of aggression. They have borne essential parts in the deeds which in the hands of their immediate perpetrators are undeniably war crimes and crimes against humanity. Yet they protest their innocence.

Our case against them has as clear a basis on the facts of this case as in the lessons of history.

They carried out orders which on the admission of many of them bit deep into the remnants of their consciences. They knew that they were doing what was wrong, but they now say "Befehl ist Befehl"—an order is an order.

All decent men find it difficult to blame others for absence of moral courage—they are only too conscious of their own failings in that direction. But there comes a point when, faced with crimes which are obvious murder or barbarity, there is a higher duty. . . . In fairness to all military tradition it should not go forth that soldiers have sheltered behind the letter of a command from facing moral problems—and deciding them, rightly or wrongly, as moral problems. Great captains are not automata to be weighed against a rubber stamp. . . . Two of the greatest names in German military history spring to one's mind: Von Clausewitz leaving the Prussian Army to serve in that of Russia; Yorck von Wartenburg making his decision of neutrality—both put what they deemed the needs of Europe and humanity above the orders of the moment. . . .

These men, of all men, knew their leader to be a callous murderer, yet for years they had met in conference after conference to sit at his feet and listen to his words. They fed his lust for power and enslavement with the best of their professional skill. While the defenseless peoples of the East, the men, women and children of Poland, of the Soviet Union, and of the Baltic states, were being deliberately slaughtered and deported into slavery to allow for German "Lebensraum," these men talked of the necessities of war. When their own cities were bombed and Germans killed, they called it murder. Only in July 1944, when Hitler's star was dimmed, did three Field Marshals and five Colonel Generals recognize that he was murdering also their own country and took action. When that star was rising in victory they had hailed it and ignored the blood-red colour of the clouds from which it rose.

Brigadier General Telford Taylor: . . . We are not concerned here with the ordinary German conscript who made up the bulk of the Wehrmacht. We are concerned entirely with professional soldiers, and with the most zealous, ambitious, and able German officers in the business. Most of them chose a military career because it was in their blood; as Manstein put it, "they considered the glory of war as something great." They slaved at it and were devoted to their profession, and if they reached the status of commander-in-chief, they were, like Manstein, proud that an army had been entrusted to them. No one became a German commander-in-chief unless he wanted to.

It is true that in time of war a professional soldier cannot resign his commission or his post of his own free will. But this does not turn the professional

officer into a conscript or make his status an involuntary one. No one becomes a professional officer without knowing in advance the obligations that will bind him in time of war. The fanatical Nazis who rushed to volunteer for the early Waffen-SS divisions or who voluntarily joined other paramilitary sections of the Party could not thereafter resign at will, but I have not heard it urged that they were conscripts or involuntary members. The members of the General Staff and High Command Group were keen professional warriors, who competed with others like themselves for the responsibilities and honors of being commanders-in-chief. They rose within the Wehrmacht just as an ambitious Party member might rise to be a Kreisleiter or Gauleiter. . . .

The German military leaders believe in war. They regard it as part of a normal, well-rounded life. Manstein told us from the witness box that they "naturally considered the glory of war as something great." The "considered opinion" of OKW in 1938 recited that:

> Despite all attempts to outlaw it, war is still a law of nature which may be challenged but not eliminated. It serves the survival of the race and state or the assurance of its historical future.
>
> This high moral purpose gives war its total character and its ethical justification.

These characteristics of the German military leaders are deep and permanent. They have been bad for the world, and bad for Germany too. Their philosophy is so perverse that they regard a lost war, and a defeated and prostrate Germany, as a glorious opportunity to start again on the same terrible cycle. . . . The German militarist will join forces with any man or government that offers fair prospect of effective support for military exploits. Men who believe in war as a way of life learn nothing from the experience of losing one. . . .

For nine months this courtroom has been a world of gas chambers, mountains of corpses, human-skin lampshades, shrunken skulls, freezing experiments, and bank vaults filled with gold teeth. It is vital to the conscience of the world that all the participants in these enormities shall be brought to justice. But these exhibits, gruesome as they are, do not lie at the heart of this case. Little will be accomplished by shaking the poisoned fruit from the tree. It is much harder to dig the tree up by the roots, but only this will in the long run do much good.

The tree which bore this fruit is German militarism. Militarism is as much the core of the Nazi Party as of the Wehrmacht itself. Militarism is not the profession of arms. Militarism is embodied in the "military-minded nation" whose leaders preach and practice conquest by force of arms, and relish war as something desirable in itself. Militarism inevitably leads to cynical and wicked disregard of the rights of others and of the very elements of civilization. Militarism destroys the moral character of the nation that practises it and, because it can be overthrown only by its own weapons, undermines the character of nations that are forced to combat it.

The wellspring of German militarism through the years had been the group of professional military leaders who have become known to the world as the "German General Staff." That is why the exposure and discrediting of this group through the declaration of criminality is far more important than the fate of the uniformed individuals in the box, or of other members of this group as individuals. . . .

The truth is spread on the record before us, and all we have to do is state the truth plainly. The German militarists joined forces with Hitler and with him created the Third Reich; with him they deliberately made a world in which might was all that mattered; with him they plunged the world into war, and spread terror and devastation over the continent of Europe. They dealt a blow at all mankind; a blow so savage and foul that the conscience of the world will reel for years to come. This was not war; it was crime. This was not soldiering; it was savagery. These things need to be said. We cannot here make history over again, but we can see that it is written true.

JUDGMENT OF THE INTERNATIONAL MILITARY TRIBUNAL

The Charter Provisions

The individual defendants are indicted under Article 6 of the Charter, which is as follows:

> Article 6. The Tribunal established by the Agreement referred to in Article 1 hereof for the trial and punishment of the major war criminals of the European

Axis countries shall have the power to try and punish persons who, acting in the interests of the European Axis countries, whether as individuals or as members of organisations, committed any of the following crimes:

The following acts, or any of them, are crimes, coming within the jurisdiction of the Tribunal for which there shall be individual responsibility:

a. Crimes against Peace: namely, planning, preparation, initiation, or waging of a war of aggression, or a war in violation of international treaties, agreements, or assurances, or participation in a common plan or conspiracy for the accomplishment of any of the foregoing.

b. War Crimes: namely, violations of the laws or customs of war. Such violations shall include, but not be limited to, murder, ill-treatment, or deportation to slave labour or for any other purpose of civilian population of or in occupied territory, murder or ill-treatment of prisoners of war or persons on the seas, killing of hostages, plunder of public or private property, wanton destruction of cities, towns, or villages, or devastation not justified by military necessity.

c. Crimes against Humanity: namely, murder, extermination, enslavement, deportation, and other inhumane acts committed against any civilian population, before or during the war, or persecutions on political, racial or religious grounds in execution of or in connection with any crime within the jurisdiction of the Tribunal, whether or not in violation of the domestic law of the country where perpetrated.

Leaders, organisers, instigators, and accomplices participating in the formulation or execution of a common plan or conspiracy to commit any of the foregoing crimes are responsible for all acts performed by any persons in execution of such plan.

These provisions are binding upon the Tribunal as the law to be applied to the case. . . .

The Law of the Charter

The jurisdiction of the Tribunal is defined in the Agreement and Charter, and the crimes coming within the jurisdiction of the Tribunal, for which there shall be individual responsibility, are set out in Article 6. The law of the Charter is decisive, and binding upon the Tribunal.

The making of the Charter was the exercise of the sovereign legislative power by the countries to which the German Reich unconditionally surrendered; and the undoubted right of these countries to legislate for the occupied territories has been recognised by the civilised world. The Charter is not an arbitrary exercise of power on the part of the victorious nations, but in the view of the Tribunal, as will be shown, it is the expression of international law existing at the time of its creation; and to that extent is itself a contribution to international law.

The Signatory Powers created this Tribunal, defined the law it was to administer, and made regulations for the proper conduct of the Trial. In doing so, they have done together what any one of them might have done singly; for it is not to be doubted that any nation has the right thus to set up special courts to administer law. With regard to the constitution of the Court, all that the defendants are entitled to ask is to receive a fair trial on the facts and law. . . .

It was submitted that international law is concerned with the action of sovereign states and provides no punishment for individuals; and further, that where the act in question is an act of state, those who carry it out are not personally responsible, but are protected by the doctrine of the sovereignty of the state. In the opinion of the Tribunal, both these submissions must be rejected. That international law imposes duties and liabilities upon individuals as well as upon states has long been recognised. In the recent case of *ex parte Quirin* before the Supreme Court of the United States, persons were charged during the war with landing in the United States for purposes of spying and sabotage. The late Chief Justice Stone, speaking for the Court, said:

> From the very beginning of its history this Court has applied the law of war as including that part of the law of nations which prescribes for the conduct of war, the status, rights, and duties of enemy nations as well as enemy individuals.

He went on to give a list of cases tried by the courts, where individual offenders were charged with offences against the laws of nations, and particularly the laws of war. Many other authorities could be cited, but enough has been said to show that individuals can be punished for violations of

international law. Crimes against international law are committed by men, not by abstract entities, and only by punishing individuals who commit such crimes can the provisions of international law be enforced. . . .

The principle of international law which, under certain circumstances, protects the representatives of a state, cannot be applied to acts which are condemned as criminal by international law. The authors of these acts cannot shelter themselves behind their official position in order to be freed from punishment in appropriate proceedings. Article 7 of the Charter expressly declares:

> The official position of defendants, whether as heads of state, or responsible officials in government departments, shall not be considered as freeing them from responsibility, or mitigating punishment.

On the other hand the very essence of the Charter is that individuals have international duties which transcend the national obligations of obedience imposed by the individual state. He who violates the laws of war cannot obtain immunity while acting in pursuance of the authority of the state, if the state in authorising action moves outside its competence under international law.

It was also submitted on behalf of most of these defendants that in doing what they did they were acting under the orders of Hitler, and therefore cannot be held responsible for the acts committed by them in carrying out these orders. The Charter specifically provides in Article 8:

> The fact that the defendant acted pursuant to order of his Government or of a superior shall not free him from responsibility, but may be considered in mitigation of punishment.

The provisions of this article are in conformity with the law of all nations. That a soldier was ordered to kill or torture in violation of the international law of war has never been recognised as a defence to such acts of brutality, though, as the Charter here provides, the order may be urged in mitigation of the punishment. The true test, which is found in varying degrees in the criminal law of most nations, is not the existence of the order, but whether moral choice was in fact possible.

REVIEW AND DISCUSSION QUESTIONS

1. Critically assess the four points that Dr. Laternser lists on page 241 in defense of the accused.
2. Laternser argues, in effect, that it is unrealistic and unfair to expect the accused to have acted other than they did. Why? Explain why you agree or disagree with his reasoning. In light of his statement and those of the prosecution, do you believe there are exculpatory factors or mitigating circumstances that should be taken into account?
3. In your view, what legal and moral duties do military officers have? What about an ordinary soldier?
4. The Tribunal asserts that international law imposes "duties and liabilities upon individuals as well as states." Explain why you agree or disagree with this.
5. Was the Tribunal established at Nuremberg legally and morally legitimate, or was it simply a matter of the Allied powers' exacting revenge?
6. The Tribunal states that the Charter under which it operates is "the expression of international law at the time of its creation" as well as "a contribution to international law." How can it be both these things?
7. The Charter defines "crimes against peace," "war crimes," and "crimes against humanity," which were the crimes the defendants were on trial for. Were these crimes in fact a recognized part of international law at the time? Explain why or why not. If they were not, was it fair to try Germany's military leaders for committing them? Are these crimes a recognized part of international law today?
8. Can international law be imposed on states like Nazi Germany that do not accept its principles?

CHAPTER 11

Law and Economics

The essays in this section concern the law and economics movement. Its basic claim is that by applying the tools of economic analysis, we are able to interpret much of the law in a way that makes it both coherent and sound. Central to this approach, as described by one of its most important defenders, Judge Richard Posner, is the goal of securing as much wealth for society as possible. By "wealth" he means people's monetary and non-monetary satisfaction based on willingness to pay. Thus, if you refuse to purchase an inexpensive safety device and your machinery causes serious injury, you may have acted unreasonably (in the sense that you have reduced total satisfaction by not taking the most cost-effective measures). According to law and economics, the law should hold you liable because doing so would encourage others to act in a more economically rational way in the future. Following Posner's discussion, Andrew Altman examines the value and limits of the law and economics approach to legal theory.

The Economic Approach to Law

Richard A. Posner

Among the recent developments in legal thinking has been a resurgence of interest in the connections between economic theory and law. Richard Posner is among the most important figures of the law and economics movement. In this selection, taken from his book *The Problems of Jurisprudence,* Posner first sets out the major tenets of his approach, including legislative bargains, the nature of wealth maximization, and the capacity of judges to promote efficiency. He shows that by viewing law as wealth-maximizing rules, we can understand not only contract and property law, but tort and criminal law as well. Richard Posner is a federal judge and professor of law at the University of Chicago Law School.

The most ambitious and probably the most influential effort in recent years to elaborate an overarching concept of justice that will both explain judicial decision making and place it on an objective basis is that of scholars working in the interdisciplinary field of "law and economics," as economic analysis of law is usually called. . . .

The basic assumption of economics that guides the version of economic analysis of law that I shall be presenting is that people are rational maximizers of their satisfactions—*all* people (with the exception of small children and the profoundly retarded) in *all* of their activities (except when under the influence of psychosis or similarly deranged through drug or alcohol abuse) that involve choice. Because this definition embraces the criminal deciding whether to commit another crime, the litigant deciding whether to settle or litigate a case, the legislator deciding whether to vote for or against a bill, the judge deciding how to cast his vote in a case, the party to a contract deciding whether to break it, the driver deciding how fast to drive, and the pedestrian deciding how boldly to cross the street, as well as the usual economic actors, such as businessmen and consumers, it is apparent that most activities either regulated by or occurring within the legal system are grist for the economic analyst's mill. It should go without saying that nonmonetary as well as monetary satisfactions enter into the individual's calculus of maximizing (indeed, money for most people is a means rather than an end) and that decisions, to be rational, need not be well thought out at the conscious level—indeed, need not be conscious at all. Recall that "rational" denotes suiting means to ends, rather than mulling things over, and that much of our knowledge is tacit.

Since my interest is in legal doctrines and institutions, it will be best to begin at the legislative (including the constitutional) level. I assume that legislators are rational maximizers of their satisfactions just like everyone else. Thus nothing they do is motivated by the public interest as such. But they want to be elected and reelected, and they need money to wage an effective campaign. This money is more likely to be forthcoming from well-organized groups than from unorganized individuals. The rational individual knows that his contribution is unlikely to make a difference; for this reason and also because voters in most elections are voting for candidates rather than policies, which further weakens the link between casting one's vote and obtaining one's preferred policy, the rational individual will have little incentive to invest time and effort in deciding whom to vote for. Only an organized group of individuals (or firms or other organizations—but these are just conduits for individuals) will be able to overcome the informational and free-rider problems that plague collective action.[1] But such a group will not organize and act effectively unless its members have much to gain or much to lose from specific policies, as tobacco farmers, for example, have much to gain from federal subsidies for growing tobacco and much to lose from the withdrawal of those subsidies. The basic tactic of an interest group is to trade the votes of its members and its financial support to candidates in exchange for an implied promise of favorable legislation. Such legislation will normally take the form of a statute transferring wealth from unorganized taxpayers (for example, consumers) to the interest group. If the target were another interest group, the legislative transfer might be effectively opposed. The unorganized are unlikely to mount effective opposition, and it is their wealth, therefore, that typically is transferred to interest groups.

On this view, a statute is a deal. . . . But because of the costs of transactions within a multi-headed legislative body, and the costs of effective communication through time, legislation does not spring full-grown from the head of the legislature; it needs interpretation and application, and this is the role of the courts. They are agents of the legislature. But to impart credibility and durability to the deals the legislature strikes with interest groups, courts must be able to resist the wishes of current legislators who want to undo their predecessors' deals yet cannot do so through repeal because the costs of passing legislation (whether original or amended) are so high, and who might therefore look to the courts for a repealing "interpretation." The impediments to legislation actually facilitate rather than retard the striking of deals, by giving interest groups some assurance that a deal struck with the legislature will not promptly be undone by repeal. An independent judiciary is one of the impediments.

Judicial independence makes the judges imperfect agents of the legislature. This is tolerable not only for the reason just mentioned but also because an independent judiciary is necessary for the resolution of ordinary disputes in a way that will encourage trade, travel, freedom of action, and other highly valued activities or conditions and will minimize the expenditure of resources on influencing governmental action. Legislators might appear to have little to gain from these widely diffused rule-of-law virtues. But if the aggregate benefits from a particular social policy are very large and no interest group's ox is gored, legislators may find it in their own interest to support the policy. Voters understand in a rough way the benefits to them of national defense, crime control, dispute settlement, and the other elements of the night watchman state, and they will not vote for legislators who refuse to provide these basic public services. It is only when those services are in place, and when (usually later) effective means of taxation and redistribution develop, that the formation of narrow interest groups and the extraction by them of transfers from unorganized groups become feasible.

The judges thus have a dual role: to interpret the interest-group deals embodied in legislation and to provide the basic public service of authoritative dispute resolution. They perform the latter function not only by deciding cases in accordance with preexisting norms, but also—especially in the Anglo-American legal system—by elaborating those norms. They fashioned the common law out of customary practices, out of ideas borrowed from statutes and from other legal systems (for example, Roman law), and out of their own conceptions of public policy. The law they created exhibits, according to the economic theory that I am expounding, a remarkable substantive consistency. It is as if the judges *wanted* to adopt the rules, procedures, and case outcomes that would maximize society's wealth.

I must pause to define "wealth maximization," a term often misunderstood. The "wealth" in "wealth maximization" refers to the sum of all tangible and intangible goods and services, weighted by prices of two sorts: offer prices (what people are willing to pay for goods they do not already own); and asking prices (what people demand to sell what they do own). If A would be willing to pay up to $100 for B's stamp collection, it is worth $100 to A. If B would be willing to sell the stamp collection for any price above $90, it is worth $90 to B. So if B sells the stamp collection to A (say for $100, but the analysis is qualitatively unaffected at any price between $90 and $100—and it is only in that range that a transaction will occur), the wealth of society will rise by $10. Before the transaction A had $100 in cash and B had a stamp collection worth $90 (a total of $190); after the transaction A has a stamp collection worth $100 and B has $100 in cash (a total of $200). The transaction will not raise measured wealth—gross national product, national income, or whatever—by $10; it will not raise it at all unless the transaction is recorded, and if it is recorded it is likely to raise measured wealth by the full $100 purchase price. But the real addition to social wealth consists of the $10 increment in *nonpecuniary* satisfaction that A derives from the purchase, compared with that of B. This shows that "wealth" in the economist's sense is not a simple monetary measure. . . .

[I]f I am given a choice between remaining in a job in which I work forty hours a week for $1,000 and switching to a job in which I would work thirty hours for $500, and I decide to make the switch, the extra ten hours of leisure must be worth at least $500 to me, yet GNP will fall when I reduce my hours of work. Suppose the extra hours of leisure are worth $600 to me, so that my full income rises from $1,000 to $1,100 when I reduce my hours. My former employer presumably is made worse off by my leaving (else why did he employ me?), but not more than $100 worse off; for if he were, he would offer to pay me a shade over $1,100 a week to stay—and I would stay. (The example abstracts from income tax.)

Wealth is *related* to money, in that a desire not backed by ability to pay has no standing—such a desire is neither an offer price nor an asking price. I may desperately desire a BMW, but if I am unwilling or unable to pay its purchase price, society's wealth would not be increased by transferring the BMW from its present owner to me. Abandon this essential constraint (an important distinction, also, between wealth maximization and utilitarianism—for I might derive greater utility from the BMW than its present owner or anyone else to whom he might sell the car), and the way is open to tolerating the crimes committed by the passionate and the avaricious against the cold and the frugal.

The common law facilitates wealth-maximizing transactions in a variety of ways. It recognizes property rights, and these facilitate exchange. It also protects property rights, through tort and criminal law. (Although today criminal law is almost entirely statutory, the basic criminal protections—for example, those against murder, assault, rape, and theft—have, as one might expect, common law origins.) Through contract law it protects the process of exchange. And it establishes procedural rules for resolving disputes in these various fields as efficiently as possible.

The illustrations given thus far of wealth-maximizing transactions have been of transactions that are voluntary in the strict sense of making everyone affected by them better off, or at least no worse off. Every transaction has been assumed to affect just two parties, each of whom has been made better off by it. Such a transaction is said to be Pareto superior [i.e., at least one person is better off, and nobody is worse off], but Pareto superiority is not a necessary condition for a transaction to be wealth maximizing. Consider an accident that inflicts a cost of $100 with a probability of .01 and that would have cost $3 to avoid. The accident is a wealth-maximizing "transaction" . . . because the expected accident cost ($1) is less than the cost of avoidance. (I am assuming risk neutrality. Risk aversion would complicate the analysis but not change it fundamentally.) It is wealth maximizing even if the victim is not compensated. The result is consistent with Learned Hand's formula, which defines negligence as the failure to take cost-justified precautions. If the only precaution that would have averted the accident is not cost-justified, the failure to take it is not negligent and the injurer will not have to compensate the victim for the costs of the accident.

If it seems artificial to speak of the accident as the transaction, consider instead the potential transaction that consists of purchasing the safety measure that would have avoided the accident. Since a potential victim would not pay $3 to avoid an expected accident cost of $1, his offer price will be less than the potential injurer's asking price and the transaction will not be wealth maximizing. But if these figures were reversed—if an expected accident cost of $3 could be averted at a cost of $1—the transaction would be wealth maximizing, and a liability rule administered in accordance with the Hand formula would give potential injures an incentive to take the measures that potential victims would pay them to take if voluntary transactions were feasible. The law would be overcoming transaction-cost obstacles to wealth-maximizing transactions—a frequent office of liability rules.

The wealth-maximizing properties of common law rules have been elucidated at considerable length in the literature of the economic analysis of law. Such doctrines as conspiracy, general average (admiralty), contributory negligence, equitable servitudes, employment at will, the standard for granting preliminary injunctions, entrapment, the contract defense of impossibility, the collateral-benefits rule, the expectation measure of damages, assumption of risk, attempt, invasion of privacy, wrongful interference with contract rights, the availability of punitive damages in some cases but not others, privilege in the law of evidence, official immunity, and the doctrine of moral consideration have been found—at least by some contributors to his literature—to conform to the dictates of wealth maximization. . . . It has even been argued that the system of precedent itself has an economic equilibrium. Precedents are created as a by-product of litigation. The greater the number of recent precedents in an area, the lower the rate of litigation will be. In particular, cases involving disputes over legal as distinct from purely factual issues will be settled. The existence of abundant, highly informative (in part because recent) precedents will enable the parties to legal disputes to form more convergent estimates of the likely outcome of a trial, and . . . if both parties agree on the outcome of trial they will settle beforehand because a trial is more costly than a settlement. But with less litigation, fewer new precedents will be produced, and the existing precedents will obsolesce as changing circumstances render them less apt and informative. So the rate of litigation will rise, producing more precedents and thereby causing the rate of litigation again to fall.

This analysis does not explain what drives judges to decide common law cases in accordance with the dictates of wealth maximization. Prosperity, however, which wealth maximization measures more sensitively than purely monetary measures such as GNP, is a relatively uncontroversial policy, and

most judges try to steer clear of controversy: their age, method of compensation, and relative weakness vis-á-vis the other branches of government make the avoidance of controversy attractive. It probably is no accident, therefore, that many common law doctrines assumed their modern form in the nineteenth century, when laissez-faire ideology, which resembles wealth maximization, had a strong hold on the Anglo-American judicial imagination. . . .

It may be objected that in assigning ideology as a cause of judicial behavior, the economist strays outside the boundaries of his discipline; but he need not rest on ideology. The economic analysis of legislation implies that fields of law left to the judges to elaborate, such as the common law fields, must be the ones in which interest-group pressures are too weak to deflect the legislature from pursuing goals that are in the general interest. Prosperity is one of these goals, and one that judges are especially well equipped to promote. The rules of the common law that they promulgate attach prices to socially undesirable conduct, whether free riding or imposing social costs without corresponding benefits.[2] By doing this the rules create incentives to avoid such conduct, and these incentives foster prosperity. In contrast, judges can, despite appearances, do little to redistribute wealth. A rule that makes it easy for poor tenants to break leases with rich landlords, for example, will induce landlords to raise rents in order to offset the costs that such a rule imposes, and tenants will bear the brunt of these higher costs. Indeed, the principal redistribution accomplished by such a rule may be from the prudent, responsible tenant, who may derive little or no benefit from having additional legal rights to use against landlords—rights that enable a tenant to avoid or postpone eviction for nonpayment of rental—to the feckless tenant. That is a capricious redistribution. Legislatures, however, have by virtue of their taxing and spending powers powerful tools for redistributing wealth. So an efficient division of labor between the legislative and judicial branches has the legislative branch concentrate on catering to interest-group demands for wealth distribution and the judicial branch on meeting the broad-based social demand for efficient rules governing safety, property, and transactions. Although there are other possible goals of judicial action besides efficiency and redistribution, many of these (various conceptions of "fairness" and "justice") are labels for wealth maximization, or for redistribution in favor of powerful interest groups; or else they are too controversial in a heterogeneous society, too ad hoc, or insufficiently developed to provide judges who desire a reputation for objectivity and disinterest with adequate grounds for their decisions.

Finally, even if judges have little commitment to efficiency, their inefficient decisions will, by definition, impose greater social costs than their efficient ones will. As a result, losers of cases decided mistakenly from an economic standpoint will have a greater incentive, on average, to press for correction through appeal, new litigation, or legislative action than losers of cases decided soundly from an economic standpoint—so there will be a steady pressure for efficient results. Moreover, cases litigated under inefficient rules tend to involve larger stakes than cases litigated under efficient rules (for the inefficient rules, by definition, generate social waste), and the larger the stakes in a dispute the likelier it is to be litigated rather than settled; so judges will have a chance to reconsider the inefficient rule.

Thus we should not be surprised to see the common law tending to become efficient, although since the incentives of judges to perform well along any dimension are weak (this is a by-product of judicial independence), we cannot expect the law ever to achieve perfect efficiency. Since wealth maximization is not only a guide in fact to common law judging but also a genuine social value and the only one judges are in a good position to promote, it provides not only the key to an accurate description of what the judges are up to but also the right benchmark for criticism and reform. If judges are failing to maximize wealth, the economic analyst of law will urge them to alter practice or doctrine accordingly. In addition, the analyst will urge—on any legislator sufficiently free of interest-group pressures to be able to legislate in the public interest—a program of enacting only legislation that conforms to the dictates of wealth maximization.

Besides generating both predictions and prescriptions, the economic approach enables the common law to be reconceived in simple, coherent terms and to be applied more objectively than

traditional lawyers would think possible. From the premise that the common law does and should seek to maximize society's wealth, the economic analyst can deduce in logical—if you will, formalist—fashion (economic theory is formulated nowadays largely in mathematical terms) the set of legal doctrines that will express and perfect the inner nature of the common law, and can compare these doctrines with the actual doctrines of common law. After translating from the economic vocabulary back into the legal one, the analyst will find that most of the actual doctrines are tolerable approximations to the implications of economic theory and so far formalistically valid. Where there are discrepancies, the path to reform is clear—yet the judge who takes the path cannot be accused of making rather than finding law, for he is merely contributing to the program of realizing the essential nature of the common law.

The project of reducing the common law—with its many separate fields, its thousands of separate doctrines, its hundreds of thousands of reported decisions—to a handful of mathematical formulas may seem quixotic, but the economic analyst can give reasons for doubting this assessment. Much of the doctrinal luxuriance of common law is seen to be superficial once the essentially economic nature of the common law is understood. A few principles, such as cost-benefit analysis, the prevention of free riding, decision under uncertainty, risk aversion, and the promotion of mutually beneficial exchanges, can explain most doctrines and decisions. Tort cases [i.e., cases involving private harms, such as negligent driving, where the injured person sues for damages—EDS.] can be translated into contract cases by recharacterizing the tort issue as finding the implied pre-accident contract that the parties would have chosen had transaction costs not been prohibitive, and contract cases can be translated into tort cases by asking what remedy if any would maximize the expected benefits of the contractual undertaking considered ex ante. The criminal's decision whether to commit a crime is no different in principle from the prosecutor's decision whether to prosecute; a plea bargain is a contract; crimes are in effect torts by insolvent defendants because if all criminals could pay the full social costs of their crimes, the task of deterring antisocial behavior could be left to tort law. Such examples suggest not only that the logic of the common law really is economics but also that the teaching of law could be simplified by exposing students to the clean and simple economic structure beneath the particolored garb of legal doctrine.

If all this seems reminiscent of Langdell, it differs fundamentally in being empirically verifiable. The ultimate test of a rule derived from economic theory is not the elegance or logicality of the derivation but the rule's effect on social wealth.

NOTES

1. A free rider is someone who derives a benefit without contributing to the cost of creating the benefit. For example, even if A and B both favor the enactment of a statute, X, each will prefer the other to invest what is necessary in getting X enacted, since the benefit of X to A or to B will be the same whether or not he contributes to the cost of obtaining it. [N]ational defense [i]s an example of an activity that would encounter severe free rider problems if provided privately.
2. Such imposition is well illustrated by acquisitive crimes: the time and money spent by the thief in trying to commit thefts and the property owner in trying to prevent them have no social product, for they are expended merely in order to bring about, or to prevent, a redistribution of wealth. Overall wealth decreases, as in the case of monopoly, discussed earlier.

REVIEW AND DISCUSSION QUESTIONS

1. Describe how Posner understands the role of the legislature, including its relationship with courts.
2. Using the example of the stamp collection, explain "wealth maximization."
3. Illustrate, using examples, how the law facilitates wealth maximization.
4. Why would a judge who made it easy for poor tenants to break leases with rich landlords not effectively benefit the poor, according to Posner?

5. Explain the formula used by Judge Learned Hand.
6. Is wealth, by itself, something of moral value? If not, does this pose a problem for Posner's theory?
7. Is Posner really a utilitarian, who uses wealth maximization as a stand-in for well-being or happiness of society? Explain.

Law and Economics: An Analysis and Critique

Andrew Altman

In this reading, taken from his book *Arguing About Law,* Andrew Altman first describes the basic approach taken by law and economics, including the idea of economic rationality and the different ways of understanding efficiency. He then discusses the strengths and weaknesses of law and economics, focusing specifically on whether or not its goal of efficient legal rules is justified and on the debate between conservatives and liberals about the economic policies that should be adopted. Andrew Altman is professor of philosophy at George Washington University.

THE ECONOMIC ANALYSIS OF LAW

Positive law is made and maintained by the decisions and actions of human beings. In light of that fact, it is reasonable to think that fields of study focusing on human behavior and institutions can contribute to the understanding of the law. Accordingly, it is unsurprising to find that such fields as sociology, psychology, and anthropology have made important contributions to the study of law. Psychologists, for example, have studied how verdicts in criminal cases are affected by whether the prosecution or the defense goes last in giving its closing statement to the jury. And anthropologists have described and contrasted the distinct ways in which different cultures handle disputes that break out among their members. But in recent years, some thinkers have proposed that the principles of economics provide the best way to describe, explain, and evaluate the rules of any system of positive law. These thinkers have established the "law and economics movement."

The law and economics movement is part of a trend that uses economic ideas, such as the concept of efficiency, to account for decisions and practices seemingly far removed from the impersonal transactions of the marketplace. For example, economic analyses have been given of dating, courtship and marriage. The underlying idea here is that the principles of economics do not apply merely to some narrowly restricted range of human behavior: they apply to human activities and practices across the board, including the law.

We should be careful to distinguish among three separate theses belonging to the law and economics approach: the descriptive, the explanatory, and the evaluative. The descriptive thesis holds that economic concepts and principles provide illuminating descriptions of legal rules. The explanatory thesis holds that economic concepts and principles help provide the best explanation for why society has the legal rules it has. And the evaluative thesis holds that economic concepts and principles provide sound criteria for evaluating legal rules and determining which ones society ought to have. . . .

The evaluative thesis of law and economics is made more specific by the claim that economic efficiency provides a criterion for evaluating the law: other things being equal, inefficient legal rules should be replaced by efficient ones, and efficient ones should be maintained. These evaluative claims about efficiency stand or fall, regardless of whether the law as a whole (or some area of it) is explained by its economic efficiency or is even efficient in the first place. The idea that the law should be efficient must be defended or criticized apart from the other claims of the law and economics movement about what the law is and why it is.

ECONOMIC RATIONALITY

In this section, we will focus on the economic concept of rationality. It is one of two concepts at the center of how the economic approach describes, explains, and evaluates human actions and institutions. The other concept is that of efficiency, which we will examine in the next main section.

RATIONAL ACTION

From the economic perspective, human action is essentially rational, and the rationality of an action is a function of its costs and benefits to the agent. *Costs* and *benefits* refer to the entire range of considerations that make something better or worse for an individual. Monetary considerations are included, but costs and benefits are not restricted to financial gains and losses. Suppose I receive an award for my work. Suppose that no money goes along with the award and that having it will not indirectly increase my income in any way (unlike, say, an Olympic gold medal). Still, the award would count as a benefit so long as it made me better off in some way, for example, by increasing my prestige among my colleagues or just by making me feel good.

According to the economic approach, each individual is the ultimate judge of what makes her better or worse off. In other words, each person's preferences determine what counts as a cost or benefit for her. Moreover, the economic approach assumes that all costs and benefits can be mathematically represented in terms of some common unit of measurement. In practice, monetary units (e.g., dollars) are used. Thus, gaining prestige among my colleagues is not itself monetary compensation, but its value to me can be represented in terms of dollars.

Rational decisions reflect the *net* benefits of the decision. In other words, they reflect the gains *minus* the losses. For economic analysts, the crucial gains and losses are the *marginal* ones. Marginal costs and benefits are the ones added to (or subtracted from) what one already has. By focusing on marginal costs and benefits, one can determine when one has maximized one's net benefits.

An Example: The Rational Athlete

Consider an athlete whose training involves sit-ups in order to increase the strength of her abdominal muscles. Suppose that we ignore all of her other aims in life and focus only on her task of strengthening her abdominal muscles. If she has done fifteen sit-ups at a certain training session, doing an extra, sixteenth sit-up will bring an added benefit. But the extra benefit will be less than the extra benefit of doing the tenth sit-up after having already done nine of them. The marginal benefit of the sixteenth sit-up is still positive, but it is not as high as the marginal benefit of the tenth.

In general, the marginal benefits of a beneficial activity decrease as the activity continues. Similarly, the marginal benefits of a good thing decrease as one obtains more of the thing: the added benefit of a loaf of bread is less to you if you already have ten loaves than if you only have one.

Benefits are maximized when one engages in a beneficial activity until the point at which its net marginal benefits drop to zero. After that point, further activity decreases one's overall benefits, while up to that point, further activity increases the benefits. Of course, even beneficial activities have costs that must be accounted for in determining their net marginal benefits, in particular, the opportunity cost of forgoing some other activity. Suppose that by doing sit-ups our athlete is forgoing another exercise, say, leg lifts, which provides the next best way of increasing the strength of her abdominal muscles. The cost of doing sit-ups is determined by how much benefit she forgoes by not doing the leg lifts. How many sit-ups would she do, assuming she is rational?

Suppose that the marginal benefit of the thirtieth sit-up is greater than that of the first leg lift, but the marginal benefit of the thirty-first sit-up is less than that of the first leg lift. She would maximize her net benefits by doing thirty sit-ups and then switching to leg lifts. How many leg lifts would she do? Assuming she is rational, she would perform each additional leg lift until the marginal benefits decrease so as to become equal to the marginal costs. Past that point, she would be detracting from her net benefits instead of maximizing them.

According to economic analysis, then, rational action maximizes one's net benefits, and maximizing one's net benefits entails acting to the point at

which marginal benefits minus marginal costs becomes zero.

Rationality and Uncertainty

In most situations in life, it is not certain what the gains and losses will be from any given choice or action. Rational choice under such conditions of uncertainty is handled using the idea of *expected* benefits (or losses). The expected benefit of a choice is simply the benefit multiplied by the probability of its occurrence. If there is a 10 percent chance (0.10) of my receiving $10, then my expected benefit is $1.

Economic analysts will concede that not all human conduct is rational in exactly the way they describe. But they will insist that their idea of rational action provides a very good approximation of much human conduct, applicable across the spectrum of human activities. Moreover, economic analysts will contend that their approach provides the most systematic and comprehensive approach to describing and explaining human conduct. It might be possible to come up with counterexamples, in which persons do not maximize their net (expected) benefits. But there is no alternative theory that does a better job than the economic approach. It may not be perfect, because humans are not always perfectly rational, but the economic approach is by far the best game in town. Or so the economic analysts of law argue.

ECONOMIC EFFICIENCY

The economic concept of efficiency concerns the costs and benefits of an action, rule, or institution to society as a whole rather than to a specific individual. The concept rests on the idea that society is simply the sum of the individuals who make it up. Accordingly, the costs and benefits to society as a whole is a function of the costs and benefits to each affected individual. But how are the costs and benefits to different individuals to be compared and weighed against one another?

Utilitarianism and Beyond

Economic thinkers developed their answer to this question by considering the problems with a utilitarian approach and formulating several alternatives to it. Let us examine the utilitarian approach and the problems it encounters.

In its classic formulation, the utilitarian approach holds that the total good to society of some act or rule is determined by the net balance of pleasure over pain produced by the act or rule. The net balance is calculated by summing up the pleasure produced for each individual (taking into account its intensity, duration, and so on) and subtracting the pain produced for each individual (also taking account of its intensity, duration, and so on). If we used this approach to define efficiency, we would say that an efficient act or rule is one that maximizes the net balance of pleasure over pain, summing up the pleasures and pain over all affected individuals.

The critical problem that economic thinkers see in this approach to defining efficiency is that it assumes one can compare the pleasures and pains of different persons and say how much more one person's pleasure (or pain) is than another's. The simple fact seems to be that there is no way to make such "interpersonal utility comparisons."

Jack may be able to say reliably that his pleasure in eating fried chicken outweighs his pain in having indigestion. Or that his pleasure in playing basketball is greater than his pleasure in playing cards. But how can anyone determine how much more (or less) Jack's pleasure in playing basketball is than Jill's pleasure in playing cards? It seems that there is no way to compare the subjective experiences of different individuals because that would demand the impossible task of getting inside their heads. Yet, if we want to know whether it is efficient for Jack and Jill to play basketball or to play cards, a utilitarian approach to efficiency demands that we make such comparisons between pleasures of Jack and those of Jill.

Pareto's Concepts of Efficiency

Pareto, an economist, developed two concepts of efficiency that circumvented the problem of interpersonal utility comparisons. The two concepts are called "Pareto optimality" and "Pareto superiority." Both concepts enable one to measure the efficiency of different ways of allocating society's resources, and both can be applied without getting

inside the minds of the affected persons and comparing their subjective experiences. Let us examine their meanings.

Suppose that all of society's resources are in a box, waiting to be allocated to its members. Suppose that X represents one specific allocation of those resources. X is Pareto-optimal if and only if there is no way to change that allocation to make at least one person better off without making anyone worse off. Conversely, X would fail to be Pareto-optimal if and only if there were some way to change the allocation so as to make at least one person better off without making anyone else worse off.

Pareto superiority involves a comparison between any two ways of allocating the same set of resources. Let X represent one way of allocating the resources and Y another. X is Pareto-superior to Y if and only if at least one person is better off with X than with Y and no one is worse off with X than with Y. And if some persons are better off with X than with Y but others are worse off with X than Y, then we cannot say that either X or Y is Pareto-superior. X and Y are noncomparable.

Both Pareto optimality and Pareto superiority dispense with interpersonal utility comparisons in the determination of efficiency. For they only require us to determine for each affected individual whether he or she would better off, worse off, or the same under various, possible allocations. And according to the economic analysts who use these Pareto criteria, it is not necessary to get inside anyone's head in order to determine whether someone is made better off or not by some change in the allocation of resources. That is because a person's *behavior* will show whether he is made better off or not. If a person would voluntarily pay (i.e., give up some valuable resource of his) in order to move from what he has with allocation X to what he would have with Y, then that person is better off with Y than with X. And if the person would pay to avoid moving from what he has with X to what he would have with Y, then he is worse off with Y than with X.

For example, suppose that Jack and Jill are the only two people involved, X consists of Jack having a cow and Jill having a horse, and Y consists of Jill having a cow and Jack having a horse. Suppose further that X is the current allocation and that Jack and Jill can get together and strike a bargain without incurring costs such as hiring lawyers and paying postage and other "transaction costs" (i.e., the costs of negotiating and making an agreement).

If Jack is willing to give Jill his cow in order to gain her horse, then he would be better off under Y. If Jill is willing to trade her horse to get Jack's cow, she too would be better off under Y. Accordingly, they will voluntarily exchange their possessions and move from X to Y. That move will be a Pareto-superior move, because it will produce an allocation that is Pareto-superior to the allocation from which they started. Whether it is Pareto-optimal will depend on whether they are willing to undertake any other exchange. If not, Y will also be Pareto-optimal.

The concepts of Pareto optimality and Pareto superiority are thus understood by economists in terms of economic exchange: each voluntary exchange produces an outcome that is Pareto-superior to the starting point (assuming that no one is affected other than the parties to the exchange), and the point at which no further voluntary exchange takes place (assuming no transaction costs) represents a Pareto-optimal distribution.

The Pareto criteria thus circumvent the problem of interpersonal utility comparisons by making the economic behavior of a person the measure of his or her well-being. Yet, certain difficulties make the criteria almost worthless in evaluating legal rules and systems.

The Limitations of Pareto's Efficiency Concepts

The first problem stems from the fact that economic exchanges typically have effects on third parties that can make them worse off. While a voluntary exchange makes the parties to the exchange better off—at least according to economic theory, it may make persons who are not parties to the exchange worse off. If I buy my groceries from Jill's store rather than Jack's, I might be making Jack worse off by increasing the profits of his competitor, who could then use the extra profits to offer better services and take away more of Jack's business.

The Pareto criteria give us no way of comparing and evaluating the outcomes of such exchanges. As soon as even one person is made worse off by some transaction, the Pareto criteria are silent on whether the transaction produces an improvement

or not. The starting point and outcome are noncomparable. Since most transactions have these third-party effects, the Pareto criteria are usually inapplicable.

Second, virtually any change in existing legal rules or social policies will produce some winners and some losers. It does not matter whether the change is a radical one (such as replacing a command economy with a market one) or a more modest one (such as lowering tax rates). In almost every case, a change in the status quo will make at least one person worse off. This means that the Pareto criteria are practically useless for judging proposed changes in legal rules and policies, for the criteria will again be silent on the question of which is better, society before the change or after.

Kaldor, Hicks, and Posner

The severe limitations of the Pareto criteria have led some economic analysts of law to develop and employ yet another standard of efficiency: the Kaldor-Hicks criterion. According to this standard, the move from one allocation, X, to another, Y, is an improvement in efficiency if and only if those who benefit from the move gain enough extra benefit so that they could fully compensate those who lose out from the move. Fully compensating a loser means giving her enough benefits so that she is as well off as she was before the move from X to Y. The Pareto demand that efficiency improvements help at least some and harm no one is replaced by the requirement that overall improvements outweigh overall harm. To that extent, the Kaldor-Hicks standard is similar to the principle of utility, which allows trade-offs between the gains of some and the losses of others.

In addition, a leading economic analyst of law, Richard Posner, has proposed a way of interpreting Kaldor-Hicks that enables it to avoid the problem of interpersonal utility comparisons. He calls his version of Kaldor-Hicks "wealth maximization," and it appears to give us the best of both worlds. We can use it to weigh the gains of some against the losses of others but without doing any interpersonal utility comparisons.

According to Posner's wealth maximization criterion, the efficient allocation of society's resources is the allocation that puts each resource in the control of the person who values it the most. And efficient legal rules are ones that serve to bring about such an allocation. But the degree to which a person values a resource is not defined in terms of the pleasure or happiness or satisfaction she would receive from possessing it. Such a definition would throw us back into the intractable problem of getting inside the heads of others. Instead, the value that a person places on possessing something is determined by her economic behavior: how much she is willing and able to pay to obtain it, if she does not already have it, or how much she asks for in order to give it up, if she does already possess it.

Posner's wealth maximization standard is the most useful of the efficiency criteria that have been developed by economic analysts of law. Accordingly, when we discuss efficiency, the idea should henceforth be understood in terms of wealth maximization.

Scarcity and Efficiency

On the economic view, efficiency is important for virtually every human society because humans generally live under conditions of scarcity. Scarcity means that the resources we need to satisfy our preferences are too few for everyone to get all that they would prefer to have.

The implication of scarcity is that some potential, preference-satisfying uses for a given resource must be denied when that resource is put to a different use. The efficient use of a resource by society means that the resource is allocated to the use(s) that has the highest value, as measured by how much people are willing to pay.

Does the law help society deal with scarcity by promoting efficiency? According to many economic analysts, the rules of common law do so: they are efficient in that they generally ensure that resources wind up in the hands of those who put them to their highest-valued use. . . .

THE EVALUATION OF LAW: SHOULD LAW MAXIMIZE WEALTH?

We now turn to the evaluative aspect of the law and economics movement: its claims about what the law should be. . . .

The Biggest Pie?

A standard claim of the law and economics movement is that efficient legal rules create the "biggest economic pie." Various interest groups may fight one another over how the slices of the pie are to be distributed, but it is in every one's interests to have legal rules that create the biggest pie. After all, the more there is to fight over, the better off everyone seems to be. For that reason, economic analysts say that efficiency is a "neutral" value that all can endorse, whatever their differences about the just or fair distribution of resources. And for that reason, economic analysts claim that efficient legal rules are, insofar as they are efficient, good legal rules.

The claim that efficient legal rules create the biggest pie seems to be conflating the idea of efficiency as wealth maximization with a very different idea of efficiency. That different idea can be called "productive efficiency," which can be explained in terms of a simple example.

Imagine a society whose sole product is widgets. The society faces a choice between mutually exclusive legal rules A and B. Suppose that the choice of rule A will lead to a higher output of widgets per work hour. We can then say that A has greater productive efficiency than B.

Legal rules can promote productive efficiency in a number of distinct ways. For example, they may create incentives for society's laborers to work harder or faster in a given period of time. Or they may create incentives for persons to come up with technological innovations that increase the output of a given period of labor time. Or they may create incentives for persons to invest in available technology that increases the output of labor.

The idea of productive efficiency is logically distinct from that of wealth maximization. For a society to have the highest possible output per work hour is one thing; for it to allocate goods and services into the hands of those who value them the most, as measured by their willingness and ability to pay, is quite different. When economic analysts of law refer to the size of the pie, sometimes they are referring to productive efficiency and sometimes to wealth maximization, and it is not always clear which one.

Moreover, it would seem misleading to refer to wealth maximization as the creation of the biggest pie possible. For the concept of wealth maximization is about who gets what slice; it is not about the overall size of the pie. Strictly speaking, it would be more accurate to claim that wealth maximization ensures that each person gets the particular slice of pie for which he or she is willing and able to pay the most.

However, Posner would argue that there is an empirical connection between wealth maximization and productive efficiency. The connection stems from the fact that the slices of the pie are economic goods that can be used to make other goods (used to "grow" the pie). The degree to which I value some good will depend, at least in part, on how productively I can use it. The more I can make with it, the more value it is to me. And when I obtain the good through market transactions, I have every incentive to use it as productively as I can, for such use maximizes its value to me.

Other Virtues of Wealth Maximization?

In Posner's view, the wealth maximization standard has many other virtues in addition to promoting productive efficiency. Among the most important is that it provides a stronger support for individual rights than utilitarianism. Consider slavery. Slaves do not have any right to their own labor: they must work for the master and cannot decide to withhold their labor or to work for someone else who is offering a better deal. Our belief that no one should be a slave rests on the idea that everyone has a right to their own labor. But what is the basis of such a right? Utilitarianism bases the "right" on the claim that slavery fails to maximize the net balance of pleasure over pain. But Posner regards that as a weak basis: there could be circumstances in which the pleasure of the slave owners and their families outweighs the pain of the slaves. He suggests that maximizing wealth, as opposed to utility, provides a stronger basis for the right to one's own labor.

The value of a slave's labor is much greater to the slave than to the owner. After all, everything the slave produces goes to the owner, with nothing to the slave. This reality creates a powerful incentive for the slave to work very unproductively. In contrast, as a free person with a right to his own labor, the individual has a strong incentive to work as productively as possible. This means that a person's labor is of greatest value to himself, and so maximizing wealth requires that the right to

such labor be given to the individual and not to anyone else.

This conclusion of Posner's may seem a bit bizarre in light of the fact that he insists that the value of something to a person is a function of how much they are willing and able to pay for it. If a slave has nothing to pay for his freedom, does it not have zero value to him, given Posner's criterion for measuring value?

Posner's answer is that the slave can borrow the amount of money that represents the value of his freedom to him. That amount would be more than the owner would be willing to pay to keep him a slave, and so wealth maximization requires that the right to a person's labor be assigned to the person himself rather than to someone else. The fact that a slave who is forced to borrow to buy his freedom would have a large debt is irrelevant in Posner's view. For wealth maximization tells us that we should assign to every person the right to freedom from the beginning. The borrowing scenario is simply a hypothetical device that helps us determine who values the slave's freedom more—the owner or the slave himself—and thus enables us to determine whether there should be a right to freedom or a right to own slaves.

Wealth Maximization and the Poor

Markets and the Work Ethic. Some critics of Posner charge that the wealth maximization standard is biased against the poor in society. They point out that a person's willingness and ability to pay for something is a function of his or her wealth. The wealthy are in a position to shell out more money for something than the indigent (or to borrow more, if they do not have enough cash on hand), and so they will end up getting much more of society's goods and services if Posner's wealth maximization standard is used. The fact that an indigent person may desire some good much more than a wealthy person or may derive much more pleasure from it is irrelevant according to Posner's standard.

Yet for Posner, one of the strengths of wealth maximization is that it does not allocate goods based simply on the intensity of a person's desire for it or the amount of pleasure she would receive from it. Such utilitarian bases for allocation flout an important moral intuition: people are not entitled to goods simply by virtue of wanting them or getting pleasure from them. Posner believes that people should have to work for what they get, and wealth maximization reinforces this work ethic. Wealth maximization supports a market society (recall the efficiency of market transactions), and in such a society the wealth of persons is generally a reflection of their productive labor. At least that is Posner's view of the matter.

The claim that wealth maximization is biased against the poor is one aspect of a broader criticism of the evaluative side of the economic approach to law. According to that criticism, reliance on notions of economic efficiency for evaluating the law is biased in favor of a right-wing political agenda. Efficiency-based recommendations, it is charged, invariably favor the sorts of solutions to social problems that political conservatives advocate.

Efficiency and the Welfare State. Posner's claim about how the market rewards the work ethic is dubious: many thinkers, including conservatives, have pointed out how success in the market is often more a matter of luck than any work ethic. But even if we reject Posner's claim, the general charge that the economic analysis of law exhibits a political bias is mistaken in a number of key respects. First, it is simply not true that efficiency-based recommendations invariably favor conservative solutions over liberal ones. Consider, for example, the question of whether government should operate a welfare program to prevent the poorest from starving or suffering similar serious deprivation. Liberals have, of course, been strong advocates of such welfare policies. Conservatives tend to be skeptical or downright hostile toward them.

Yet, Posner quite plausibly points out that a welfare system could well be justified on the basis of economic efficiency. Private charity is likely to prove insufficient, and starving persons will not generally die quietly in the street. They will commit crimes in order to feed themselves and their children. It may well be more efficient economically to establish welfare programs to ensure that such persons are fed than to leave them no choice but to feed themselves through criminal activity. That is because crime creates very heavy social costs: people spend a great deal to avoid being crime victims

(buying alarm systems, not going out at night, etc.), and society spends a great deal investigating and prosecuting crimes and in imprisoning offenders. Running a welfare system may not be cheap, but it may well be cheaper for society than paying the increased costs of crime that would ensue were there no welfare system.

Another efficiency reason for some kind of a welfare program is rooted in the altruistic desires of many citizens. These citizens have a preference that their fellow citizens not suffer from drastic deprivation. If enough people who are well off have such a preference, then a welfare program might be called for on efficiency grounds.

It may be asked why efficiency cannot be served by having altruistic citizens donate to private charity. What is the need for the liberal, government imposed solution? Posner points out that there is an efficiency-based justification for reliance on government in this context. Although a well-off person may really want to see all the poor fed, the donations to private charity of others may make her contribution unnecessary. And her contribution, by itself, will not determine whether there are enough total contributions so that all the poor can be fed. Thus, a rational incentive exists for the well-off person to hold back from giving to private charity—in the hopes that others will give and make her contribution unnecessary—so that she can be a "free rider." As a free rider, she would get what she wants—all the poor fed—but not have to pay for it. But since each well-off person would have a rational incentive to refrain from donating so she could be a free rider, the result is likely to be a lower level of actual charitable contributions than would be efficient. And as Posner points out, government action solves this free-rider problem by financing a welfare program out of taxes that the well-off person has no choice but to pay.

There are many social problems in which a free-rider analysis can lead to liberal solutions. Environmental problems are rife with free-rider considerations, and many proponents of law and economics argue that strict environment laws of the sort typically favored by political liberals are needed to promote efficiency. This is not to say that there is unanimity among the proponents of law and economics on environmental issues or any other politically controversial topic. But that very point takes us to the nub of the problem with the claim that law and economics has an inherent conservative bias: many proponents of law and economics use its ideas to advocate liberal solutions to policy issues.

POLITICAL DISAGREEMENT IN LAW AND ECONOMICS

What Kind of Market?

The views that economic analysts of law have on current issues span the political spectrum from liberal to conservative. There is general agreement among them that Kaldor-Hicks efficiency is a politically neutral value that society should take account of in developing its law. And the free market is virtually unanimously endorsed as the main engine of the economics system. But there are serious disagreements as well that lead to very different conclusions about what kinds of legal policies society should adopt.

Roughly, we can distinguish between a conservative and a liberal side to the law and economics movement. The two sides disagree over the extent to which the market should be left to operate free of legal interference and regulation. The liberal side advocates greater legal regulation and interference than does the conservative. The disagreement is rooted in conflicting conceptions of the efficiency and morality of unregulated market outcomes.

Conservatives versus Liberals

Those on the conservative side of the law and economics movement tend to think that the market almost always produces efficient outcomes and so is only rarely in need of legal correction in order to achieve efficiency. They do not reject the existence of transaction costs and other factors that interfere to some extent with the efficiency of market operations. But they judge that such interferences are usually minimal and that legal efforts to alter market outcomes usually create more inefficiency than they correct.

In contrast, the liberal side sees a greater need for legal correction of the market. It holds that transaction costs, free-rider problems, and other factors that

interfere with market efficiency are more extensive than the conservative side admits. And it believes that legal regulation or circumvention of the market can effectively counteract such costs and problems and yield efficient outcomes for society.

In addition, those on the liberal side tend to subscribe to theories of justice that claim the priority of justice over efficiency and that require substantial redistribution from the outcomes produced by the market, even when the market is operating efficiently. For example, their theories of justice generally hold that government must provide significant assistance to the indigent and significantly reduce the disparities in wealth and income that result from the operation of market forces.

In contrast, those on the conservative side either deny that justice is a value that can take priority over efficiency or else hold a theory of justice according to which market outcomes are essentially just. Thus, the conservatives argue that, when transaction costs do interfere with the efficiency of market exchange, the law should generally "mimic the market." In other words, the law should produce the same outcome the market would have produced had transaction costs not interfered.

Conservatives who deny that justice takes priority over efficiency tend to be skeptical of all value judgments other than those about efficiency and to adopt a "deflationary" attitude toward such judgments. They regard judgments about justice and other moral values as no more than expressions of personal preference. Such subjective moral preferences can be incorporated along with all the other preferences people have in order to decide what is most efficient, that is, what maximizes the satisfaction of preferences. But these law and economics conservatives doubt that justice and other moral values can provide legitimate constraints on the pursuit of efficiency.

Notice that this deflationary attitude toward values other than efficiency sharply distinguishes these law-and-economics conservatives from many of the political groups in society that are regarded as conservative.

For example, consider fundamentalist religious groups that favor state-sponsored prayer in school, oppose homosexual rights, reject abortion, and otherwise argue that moral values must be enforced by the government. Such groups do not conceive of the moral values they advocate as personal preferences to be cranked into the efficiency equation along with the differing personal preferences of their political opponents. Rather, they think of their moral values as the true ones and as disqualifying from consideration any preferences that conflict with those values. Such persons might consistently adopt the other elements of conservative law and economics: the efficiency of markets and so on. But they would have to judge that, where market transactions would violate their moral values, the market must yield.

Some law-and-economics conservatives reject the deflationary account of justice and morality. They believe that justice is the primary value to which social institutions must conform. But they wind up with the same conclusion that the deflationary conservatives endorse: the free market should be largely left alone to operate without legal interference or regulation. This is because they hold to a theory of justice that says that market outcomes are largely just. Liberals, in turn, will counter with a view that questions the justice of market outcomes.

REVIEW AND DISCUSSION QUESTIONS

1. Explain what is meant by economic efficiency and by the economic conception of rational action.
2. Discuss the different ways of understanding economic efficiency, including the views of Pareto, Kaldor and Hicks, and Posner.
3. Explain and assess the claim that efficient legal rules create the biggest pie.
4. What is Posner's approach to slavery? To welfare programs?
5. Assess the charge that the wealth maximization standard is biased against the poor.
6. Contrast the conservative and the liberal sides of the law-and-economics movement.
7. What are the advantages and disadvantages of using wealth maximization as a test for legal rules?

PART FOUR

Philosophical Issues in Criminal Law

CHAPTER 12

The Justification of Punishment

Punishment involves inflicting harm on a person; that is, it involves doing something to someone that would normally be wrong to do to another person (such as taking away his or her property or freedom). Because of this, the practice of punishment requires moral justification. Punishment is often defended because of its usefulness: It deters potential criminals and protects society by rehabilitating the offenders or by removing them from society. Some contend that punishment cannot be morally justified on the basis of social utility but only on the ground that the criminal *deserves* to be punished as a result of his or her wrongdoing. The essays of this section explore these and other conflicting approaches to punishment. Richard B. Brandt and Herbert Morris present, respectively, the utilitarian and retributivist approaches to punishment. Jean Hampton argues for the moral education theory of punishment as an alternative to both, while Randy Barnett defends restitution as replacement for punishment.

Who Should Be Punished?

The Case of the Dog Provetie

In this case from 1595, a dog named Provetie was accused of biting a child in the hand, a wound that led to the child's death. The court finds Provetie guilty, condemning him to be hung on the gallows like other murderers. Although from a modern perspective the case seems bizarre, trials of animals were recorded throughout the Middle Ages. They raise a basic philosophical question: Who should be punished and why?

Lot Huygens Gael, Schout of the Town of Leiden, prosecutor on behalf of his lordship [the Count of Holland] in criminal matters, accuses in the open Court of the Schepenes of the Town of Leiden the dog of Jan Jansse van der Poel, named Provetie, or by whatsoever other name he may be called, now a prisoner, and says that he, the said Provetie, did not scruple on Sunday last, being the 5th of May, 1595,

to bite the child of Jan Jacobsz van der Poel, which child was then playing at his uncle's house and had a piece of meat in his hand, and the said Provetie snapping at it did bite the said child and thus inflicted a wound in the second finger of the right hand, going through the skin to the flesh in such manner that the blood flowed therefrom, and the child a few days after died in consequence of fright, for which cause the prosecutor apprehended the said Provetie, all of which appears from the prisoner's own confession, made by him without torture or being put in irons. . . .

Sentence: The Schepenen of Leiden, having seen the claim and conclusion made and taken by Lot Huygens Gael, Schout of this town, against and to the charge of the dog of Jan Jansse van der Poel, named Provetie, or by whatsoever other name or surname he may be known, the prisoner being present, having seen, moreover, the information obtained by the prosecutor for the purpose, besides the prisoner's own confession made without torture or being placed in irons, doing justice in the name of, etc., have condemned and hereby do condemn him to be led and taken to the plain of Gravesteijn in this town, where evildoers are customarily punished, and that he be there hanged by the executioner to the gallows with a rope until death ensues, that further his dead body be dragged on a hurdle to the gallows-field, and that he there remain hanging to the gallows, to the deterring of other dogs and to all as an example; moreover, they declare all his goods, should he have any, to be confiscated and forfeited for the benefit of the countship.

This done in the open court, all schepenen being present, the 15th May, 1595.

REVIEW AND DISCUSSION QUESTIONS

1. Owners of dogs frequently discipline their pets for bad behavior, and many believe that their pets know when they have misbehaved. If so, then why do we find it odd for this court to have punished a dog?
2. Did Provetie deserve to die? Did he deserve to be punished in some other way for biting the child's hand? Was Provetie responsible for his actions? Are we responsible for ours?
3. The court orders Provetie's dead body to be displayed "to the deterring of other dogs and to all as an example." Can dogs be deterred from bad behavior by the punishing of other dogs? Can human beings be deterred by punishing dogs?
4. Could one justify executing Provetie on grounds of social protection? What would a utilitarian say about punishing dogs?
5. When the court refers to "the prisoner's own confession," what could it have had in mind?
6. What exactly is it about human beings that makes us believe that punishing people for their conduct is sometimes appropriate or justified, whereas punishing dogs is inappropriate?

The Utilitarian Theory of Criminal Punishment

Richard B. Brandt

Utilitarianism maintains that actions, institutions, and social policies are morally justified insofar as they bring about more net social good (or happiness or well-being) than any alternative action, institution, or policy. Beginning with the views of Jeremy Bentham, one of the founders of utilitarianism, Richard B. Brandt, who was professor of philosophy at the University of Michigan, explains the utilitarian approach to criminal justice and defends it against several common criticisms. He concludes by raising some difficulties for nonutilitarian or retributive theories of punishment.

What is meant by an "examination of the ethical foundations of the institution and principles of criminal justice"? The job of such an examination is *not* to provide a moral blessing for the status quo, for the system of criminal justice as it actually is in the United States. . . . Rather, it is to identify the more important valid ethical principles that are relevant to the institution of criminal justice and to furnish a model of their use in criticism or justification of important features of this institution.

The broad questions to be kept in the forefront of discussion are the following: (1) What justifies anyone in inflicting pain or loss on an individual on account of his past acts? (2) Is there a valid general principle about the punishments proper for various acts? (Possibly there should be no close connection between offense and penalty; perhaps punishment should be suited to the individual needs of the criminal, and not to his crime.) (3) What kinds of defense should excuse from punishment? An answer to these questions would comprise prescriptions for the broad outlines of an ideal system of criminal justice.

In our [earlier] discussion of "distributive justice," we decided that "to act unjustly" means the same as "to treat unequally, in some matter that involves the distribution of things that are good or bad, except as the inequality is required by moral considerations (principles) with substantial weight in the circumstances." If this definition of "act unjustly" is correct, then there are two distinct ways in which there can be injustice in the treatment of criminals. First, criminals are *punished* whereas noncriminals are not. Punishment, however, is *unequal* treatment, in a matter that involves distribution of things good or bad. Therefore, if punishment is to be just, it must be shown that the unequal treatment is required by moral principles of weight. Thus, one thing that must be done in order to show that the practice of punishing criminals is not unjust, is to show that there are moral principles that require it. But second, the *procedures of applying* the principles directing unequal treatment for criminals may themselves operate unequally. One man gets a "fair" trial and another does not. There can be inequality in the chances given people to escape the application of legal sanctions in their case. Part of treating people "justly," then, is providing legal devices so that everyone has an equal hearing: scrupulous adherence to the rules of evidence, opportunity for appeal to higher courts for remedy of deviation from standard rules in the lower courts, and so on. We shall not here consider details about how legal institutions should be devised in order to secure equal application of the law; that is a specialized inquiry that departs too far from the main problems of ethical principle. It is a part of "justice," however. Indeed, we may view "criminal justice" as having two main aspects: just laws for the punishment of offenders and procedures insuring just application of these laws by the courts and other judicial machinery.

The existence of just laws directing certain punishments for certain offenses, then, is not the whole of justice for the criminal, but we shall concentrate on identifying such laws. . . .

THE UTILITARIAN THEORY

Historically there has been a cleavage of opinion about the kind of general ethical principles required for coherence with our concrete justified beliefs about criminal justice. Many writers have thought that a utilitarian principle is adequate. Others have thought that some nonutilitarian principle, or more than one, is necessary. Most of the latter writers (formalists) have espoused some form of *retributive* principle—that is, a principle roughly to the effect that a wrongdoer should be punished approximately in correspondence with either the moral reprehensibility of his offense or with the magnitude of his breach or of the public harm he commits.

It is convenient to begin with the utilitarian theory. . . . The essence of the rule-utilitarian theory, we recall, is that our actions, whether legislative or otherwise, should be guided by a set of prescriptions, the conscientious following of which by all would have maximum net expectable utility. As a result, the utilitarian is not, just as such, committed to any particular view about how anti-social behavior should be treated by society—or even to the view that society should do anything at all about immoral conduct. It is only the utilitarian principle *combined* with statements about the kind of laws and practices which

will maximize expectable utility that has such consequences. Therefore, utilitarians are free to differ from one another about the character of an ideal system of criminal justice; some utilitarians think that the system prevalent in Great Britain and the United States essentially corresponds to the ideal, but others think that the only system that can be justified is markedly different from the actual systems in these Western countries. We shall concentrate our discussion, however, on the more traditional line of utilitarian thought which holds that roughly the actual system of criminal law, say in the United States, is morally justifiable, and we shall follow roughly the classic exposition of the reasoning given by Jeremy Bentham[1]—but modifying this freely when we feel amendment is called for. At the end of the chapter we shall look briefly at a different view.

Traditional utilitarian thinking about criminal justice has found the rationale of the practice, in the United States, for example, in three main facts. (Those who disagree think the first two of these "facts" happen not to be the case.) (1) People who are tempted to misbehave, to trample on the rights of others, to sacrifice public welfare for personal gain, can usually be deterred from misconduct by fear of punishment, such as death, imprisonment, or fine. (2) Imprisonment or fine will teach malefactors a lesson; their characters may be improved, and at any rate a personal experience of punishment will make them less likely to misbehave again. (3) Imprisonment will certainly have the result of physically preventing past malefactors from misbehaving, during the period of their incarceration.

In view of these suppositions, traditional utilitarian thinking has concluded that having laws forbidding certain kinds of behavior on pain of punishment, and having machinery for the fair enforcement of these laws, is justified by the fact that it maximizes expectable utility. Misconduct is not to be punished just for its own sake; malefactors must be punished for their past acts, according to law, as a way of maximizing expectable utility.

The utilitarian principle, of course, has implications for decisions about the severity of punishment to be administered. Punishment is itself an evil, and hence should be avoided where this is consistent with the public good. Punishment should have precisely such a degree of severity (not more or less) that the probable disutility of greater severity just balances the probable gain in utility (less crime because of the more serious threat). The cost, in other words, should be counted along with the value of what is bought; and we should buy protection up to the point where the cost is greater than the protection is worth. How severe will such punishment be? Jeremy Bentham had many sensible things to say about this. Punishment, he said, must be severe enough so that it is to no one's advantage to commit an offense even if he receives the punishment; a fine of $10 for bank robbery would give no security at all. Further, since many criminals will be undetected, we must make the penalty heavy enough in comparison with the prospective gain from crime, that a prospective criminal will consider the risk hardly worth it, even considering that it is not certain he will be punished at all. Again, the more serious offenses should carry the heavier penalties, not only because the greater disutility justifies the use of heavier penalties in order to prevent them, but also because criminals should be motivated to commit a less serious rather than a more serious offense. Bentham thought the prescribed penalties should allow for some variation at the discretion of the judge, so that the actual suffering caused should roughly be the same in all cases; thus, a heavier fine will be imposed on a rich man than on a poor man.

Bentham also argued that the goal of maximum utility requires that certain facts should *excuse* from culpability, for the reason that punishment in such cases "must be inefficacious." He listed as such (1) the fact that the relevant law was passed only after the act of the accused, (2) that the law had not been made public, (3) that the criminal was an infant, insane, or was intoxicated, (4) that the crime was done under physical compulsion, (5) that the agent was ignorant of the probable consequences of his act or was acting on the basis of an innocent misapprehension of the facts, such that the act the agent thought he was performing was a lawful one, and (6) that the motivation to commit the offense was so strong that no threat of law could prevent the crime. Bentham also thought that punishment should be remitted if the crime was a collective one and the number of the guilty so large that great suffering

would be caused by its imposition, or if the offender held an important post and his services were important for the public, or if the public or foreign powers would be offended by the punishment; but we shall ignore this part of his view.

Bentham's account of the logic of legal "defenses" needs amendment. What he should have argued is that *not* punishing in certain types of cases (cases where such defenses as those just indicated can be offered) reduces the amount of suffering imposed by law and the insecurity of everybody, and that failure to impose punishment in these types of cases will cause only a negligible increase in the incidence of crime.

How satisfactory is this theory of criminal justice? Does it have any implications that are far from being acceptable when compared with concrete justified convictions about what practices are morally right?[2]

Many criminologists would argue that Bentham was mistaken in his facts: The deterrence value of threat of punishment, they say, is much less than he imagined, and criminals are seldom reformed by spending time in prison. If these contentions are correct, then the ideal rules for society's treatment of malefactors are very different from what Bentham thought, and from what actual practice is today in the United States. To say all this, however, is not to show that the utilitarian *principle* is incorrect. . . . Utilitarian theory might still be correct, but its implications would be different from what Bentham thought. . . .

The whole utilitarian approach, however, has been criticized on the grounds that it ought not in consistency to approve of *any* excuses from criminal liability.[3] Or at least, it should do so only after careful empirical inquiries. It is not obvious, it is argued, that we increase net expectable utility by permitting such defenses. At the least, the utilitarian is committed to defend the concept of "strict liability." Why? Because we could get a more strongly deterrent effect if everyone knew that *all behavior* of a certain sort would be punished, irrespective of mistaken supposals of fact, compulsion, and so on. The critics admit that knowledge that all behavior of a certain sort will be punished will hardly deter from crime the insane, persons acting under compulsion, persons acting under erroneous beliefs about facts, and others, but, as Professor Hart points out, it does not follow from this that general knowledge that certain acts will always be punished will not be salutary.

The utilitarian, however, has a solid defense against charges of this sort. We must bear in mind (as the critics do not) that the utilitarian principle, *taken by itself, implies nothing whatever* about whether a system of law should excuse persons on the basis of certain defenses. What the utilitarian does say is that, when we *combine* the principle of utilitarianism with *true* propositions about a certain thing or situation, then we shall come out with true statements about obligations. . . . In fact the utilitarian can properly claim that we do have excellent reason for believing that the general public would be no better motivated to avoid criminal offenses than it now is, if the insane and others were also punished along with intentional wrongdoers. Indeed, he may reasonably claim that the example of punishment of these individuals could only have a hardening effect—like public executions. Furthermore, the utilitarian can point out that abolition of the standard exculpating excuses would lead to serious insecurity. Imagine the pleasure of driving an automobile if one knew one could be executed for running down a child whom it was absolutely impossible to avoid striking! One certainly does not maximize expectable utility by eliminating the traditional excuses. In general, then, the utilitarian theory is not threatened by its implications about exculpating excuses.

It might also be objected against utilitarianism that it cannot recognize the validity of *mitigating* excuses. Would not consequences be better if the distinction between premeditated and impulsive acts were abolished? The utilitarian can reply that people who commit impulsive crimes, in the heat of anger, do not give thought to legal penalties; they would not be deterred by a stricter law. Moreover, such a person is unlikely to repeat his crime, so that a mild sentence saves an essentially good man for society.[4]

Sometimes it is objected to utilitarianism that it must view imprisonment for crime as morally no different from quarantine. This, it is said, shows that the utilitarian theory must be mistaken, since actually there is a vast moral difference between being quarantined and being imprisoned for crime. *Why* is it supposed utilitarian theory must view imprisonment as a kind of quarantine? The answer is that utilitarianism looks to the

future; the treatment it prescribes for individuals is treatment with an eye to maximizing net expectable utility. The leper is quarantined because otherwise he will expose others to disease. The criminal is imprisoned because otherwise he, or others who are not deterred by the threat of punishment, will expose the public to crime. Both the convicted criminal and the leper are making contributions to the public good. So, quarantine and imprisonment are essentially personal sacrifices for the public welfare, if we think of punishment as the utilitarian does. But in fact, the argument goes on, we feel there is a vast difference. The public is obligated to do what is possible to make the leper comfortable, to make his necessary sacrifice as easy for him and his family as possible. But we feel no obligation to make imprisonment as comfortable as possible.

Again the utilitarian has a reply. He can say that people cannot help contracting leprosy, but they can avoid committing crimes—and the very discomforts and harshness of prison life are deterring factors. If prison life were made attractive, there might be more criminals—not to mention the indolent who would commit a crime in order to enjoy the benefits of public support. Furthermore, the utilitarian can say, why should we feel that we "ought to make it up to" a quarantined leper? At least partly because it is useful to encourage willingness to make such sacrifices. But we do not at all wish to encourage the criminal to make his "sacrifice"; rather, we wish him not to commit his crimes. There is all the difference between the kind of treatment justified on utilitarian grounds for a person who may have to make a sacrifice for the public welfare through no fault of his own, and for a person who is required to make a sacrifice because he has selfishly and deliberately trampled on the rights of others, in clear view of the fact that if he is apprehended society must make an example of him. There are all sorts of utilitarian reasons for being kindly to persons of the former type, and stern with people of the latter type.

Another popular objection to the utilitarian theory is that the utilitarian must approve of prosecutors or judges occasionally withholding evidence known to them, for the sake of convicting an innocent man, if the public welfare really is served by so doing. Critics of the theory would not deny that there *can* be circumstances where the dangers are so severe that such action is called for; they only say that utilitarianism calls for it all too frequently. Is this criticism justified? Clearly, the utilitarian is not committed to advocating that a provision should be written into the *law* so as to permit punishment of persons for crimes they did not commit if to do so would serve the public good. Any such provision would be a shattering blow to public confidence and security. The question is only whether there should be an informal moral rule to the same effect, for the guidance of judges and prosecutors. Will the rule-utilitarian necessarily be committed to far too sweeping a moral rule on this point? We must recall that he is not in the position of the act-utilitarian, who must say that an innocent man must be punished if in *his particular case* the public welfare would be served by his punishment. The rule-utilitarian rather asserts only that an innocent man should be punished if he falls within a class of cases such that net expectable utility is maximized if *all* members of the class are punished, taking into account the possible disastrous effects on public confidence if it is generally known that judges and prosecutors are guided by such a rule. . . . When we take these considerations into account, it is *not* obvious that the rule-utilitarian is committed to action that we are justifiably convinced is immoral.[5] . . .

Everything considered, the utilitarian theory seems to be in much less dire distress, in respect of its implications for criminal justice, than has sometimes been supposed. It does not seem possible to show that in any important way its implications are clearly in conflict with our valid convictions about what is right. The worst that can be said is that utilitarian theory does not in a clear-cut way definitely require us to espouse some practices we are inclined to espouse. But to this the utilitarian may make two replies. First, that there is reason to think our ordinary convictions about punishment for crime ought to be thoroughly re-examined in important respects. Second, the utilitarian may reply that if we consider our convictions about the punishments we should administer *as a parent*—and this is the point where our moral opinions are least likely to be affected by the sheer weight of tradition—we shall find that we think according to the principles of rule-utilitarianism. Parents do regard their

punishment of their children as justified only in view of the future good of the child, and in order to make life in the home tolerable and in order to distribute jobs and sacrifices equally.

THE RETRIBUTIVE THEORY OF CRIMINAL JUSTICE

. . . [The] "retributive theory" . . . asserts that it is a basic principle of ethics roughly that pain or loss should be caused to persons who have done wrong, with a severity corresponding with the moral gravity of their deed—and of course the "gravity" of the deed not being defined to accord exactly with the utilitarian theory about how severely wrongdoers should be made to suffer. In saying that such a principle is a "basic" principle of ethics, proponents of the retributive theory deny the possibility of deriving this principle from any principle directing to do good, that is, from any kind of utilitarian principle. . . .

The traditional retributive principle is perhaps best stated today in a way suggested by Ross's formalist system, somewhat as follows: "It is prima facie obligatory for society to cause pain or loss to every person who commits a morally objectionable act to an extent corresponding with the moral gravity of his offense." We can assume that other considerations, such as the obligation to avoid general insecurity, will require that punishment be imposed only for infractions of properly publicized laws, by specially authorized persons, and after a trial according to procedures selected in order to guarantee a fair application of the law.

Should we accept the retributive principle as a basic "axiom" about moral obligation (or else the assertion that it is intrinsically better for offenders to be punished than to go unpunished)? Various considerations suggest that we should answer this question *negatively*:

1. Our ethical theory is *simpler* without this principle, and therefore it should be rejected unless it enables us to deduce, as theorems (when we combine it with true factual premises), ethical principles which are valid, and which cannot be deduced without it. But since our discussion of the rule-utilitarian theory of punishment has not disclosed any major objection to that theory, there is no reason to complicate our theory by adding a retributive principle.

2. The retributive principle asserts in effect that a principle aim of the law is to punish moral guilt. But if so, then it ought to punish merely *attempted* crimes as severely as successful crimes, and since an attempt is a case of setting oneself to commit a crime, it is as much a deliberate deviation from subjective obligation as the successful commission of a crime. Assuming that this implication is incorrect, clearly the retributive principle alone will not do as a principle guiding legislative practice.

3. According to retributivism, laws should be so framed that no one will be punished, no matter what he does, if he is morally blameless. This is objectionable. It is of great importance that the law be able to set up standards of conduct, and require all to conform, whether or not they are convinced of the desirability of the standards. The law must be in a position to demand certain conduct from individuals, say in the Defense Department, whose conscientious deliberations might lead them to betray secrets essential to the national defense. Again, the law must be in a position to ban some practice like polygamy, irrespective of the value judgments of any persons. Therefore, we must again say that the retributive principle cannot be the only principle guiding the framing of law and judicial practice.

NOTES

1. In *Principles of Morals and Legislation.*
2. Act-utilitarians face some special problems. For instance, if I am an act-utilitarian and serve on a jury, I shall work to get a verdict that will do the most good, irrespective of the charges of the judge, and of any oath I may have taken to give a reasonable answer to certain questions on the basis of the evidence presented—unless I think my doing so will have indirect effects on the institution of the jury, public confidence in it, and so on. This is certainly not what we think a juror should do. Of course, neither a juror nor a judge can escape his prima facie obligation to do what good he can; this obligation is present in some form in every theory. The act-utilitarian, however, makes this the whole of one's responsibility.

3. See H.L.A. Hart, "Legal Responsibility and Excuses," in Sidney Hook (ed.), *Determinism and Freedom* (New York: New York University Press, 1958), pp. 81–104; and David Braybrooke, "Professor Stevenson, Voltaire, and the Case of Admiral Byng," *Journal of Philosophy*, LIII (1956), 787–96.
4. The utilitarian must admit that the same thing is true for many deliberate murders and probably he should also admit that some people who commit a crime in the heat of anger would have found time to think had they known that a grave penalty awaited them.
5. In any case, a tenable theory of punishment must approve of punishing persons who are *morally* blameless. Suppose someone commits treason for moral reasons. We may have to say that his deed is not reprehensible at all, and might even (considering the risk he took for his principles) be morally admirable. Yet we think such persons must be punished no matter what their motives; people cannot be permitted to take the law into their own hands.

REVIEW AND DISCUSSION QUESTIONS

1. What are the two distinct ways in which there can be injustice in the treatment of criminals?
2. From a utilitarian perspective, what reasons are there for punishment? What are the implications of the utilitarian approach for the severity of punishment? for permitting exculpatory or mitigating excuses?
3. How satisfactory is Brandt's response to the criticism that utilitarianism must approve of prosecutors or judges occasionally convicting innocent people, if the public welfare would be served by doing so?
4. Assess Brandt's three reasons for rejecting the retributivist principle.
5. Are there possible objections to the utilitarian theory of punishment that Brandt has overlooked or failed to deal with adequately?
6. Assess capital punishment from a utilitarian viewpoint. What considerations can be raised for and against it? When, if ever, and for what crimes and under what circumstances would a utilitarian favor the death penalty?

Persons and Punishment

Herbert Morris

In this essay, professor of law Herbert Morris upholds the strange-sounding idea that we have a right to punishment. He does this by contrasting two systems of social control. The first is a system of just punishment. Where the rules are designed to benefit everyone, punishment prevents those who violate the rules from gaining an unfair advantage over those who adhere to them, and the offender can be seen as owing a debt to society. The second system is one that views criminal behavior as a kind of disease and sees the criminal as in need of therapeutic treatment. Because Morris finds such a therapy system morally objectionable, he concludes that we have not just a right to a system of punishment (rather than therapy) but, once such a system is in existence, a right to be punished.

> They acted and looked . . . at us, and around in our house, in a way that had about it the feeling—at least for me—that we were not people. In their eyesight we were just things, that was all.
>
> —Malcolm X

> We have no right to treat a man like a dog.
>
> —Governor Maddox of Georgia

Alfredo Traps in Dürrenmatt's tale discovers that he has brought off, all by himself, a murder involving considerable ingenuity. The mock prosecutor in the tale demands the death penalty "as reward for a crime that merits admiration, astonishment, and respect." Traps is deeply moved; indeed, he is exhilarated, and the whole of his life becomes more heroic, and, ironically,

more precious. His defense attorney proceeds to argue that Traps was not only innocent but incapable of guilt, "a victim of the age." This defense Traps disavows with indignation and anger. He makes claim to the murder as his and demands the prescribed punishment—death.

The themes to be found in this macabre tale do not often find their way into philosophical discussions of punishment. These discussions deal with large and significant questions of whether or not we ever have the right to punish, and if we do, under what conditions, to what degree, and in what manner. There is a tradition, of course, not notable for its present vitality, that is closely linked with motifs in Dürrenmatt's tale of crime and punishment. Its adherents have urged that justice requires a person be punished if he is guilty. Sometimes—though rarely—these philosophers have expressed themselves in terms of the criminal's *right to be punished.* Reaction to the claim that there is such a right has been astonishment combined, perhaps, with a touch of contempt for the perversity of the suggestion. A strange right that no one would ever wish to claim! With that flourish the subject is buried and the right disposed of. In this paper the subject is resurrected.

My aim is to argue . . . that we have a right to punishment. . . .

When someone claims that there is a right to be free, we can easily imagine situations in which the right is infringed and easily imagine situations in which there is a point to asserting or claiming the right. With the right to be punished, matters are otherwise. The immediate reaction to the claim that there is such a right is puzzlement. And the reasons for this are apparent. People do not normally value pain and suffering. Punishment is associated with pain and suffering. When we think about punishment we naturally think of the strong desire most persons have to avoid it, to accept, for example, acquittal of a criminal charge with relief and eagerly, if convicted, to hope for pardon or probation. Adding, of course, to the paradoxical character of the claim of such a right is difficulty in imagining circumstances in which it would be denied one. When would one rightly demand punishment and meet with any threat of the claim being denied?

So our first task is to see when the claim of such a right would have a point. I want to approach this task by setting out two complex types of institutions both of which are designed to maintain some degree of social control. In the one a central concept is punishment for wrongdoing and in the other the central concepts are control of dangerous individuals and treatment of disease.

Let us first turn attention to the institutions in which punishment is involved. The institutions I describe will resemble those we ordinarily think of as institutions of punishment; they will have, however, additional features we associate with a system of just punishment.

Let us suppose that men are constituted roughly as they now are, with a rough equivalence in strength and abilities, a capacity to be injured by each other and to make judgments that such injury is undesirable, a limited strength of will, and a capacity to reason and to conform conduct to rules. Applying to the conduct of these men are a group of rules, ones I shall label "primary," which closely resemble the core rules of our criminal law, rules that prohibit violence and deception and compliance with which provides benefits for all persons. These benefits consist in noninterference by others with what each person values, such matters as continuance of life and bodily security. The rules define a sphere for each person, then, which is immune from interference by others. Making possible this mutual benefit is the assumption by individuals of a burden. The burden consists in the exercise of self-restraint by individuals over inclinations that would, if satisfied, directly interfere or create a substantial risk of interference with others in proscribed ways. If a person fails to exercise self-restraint even though he might have and gives in to such inclinations, he renounces a burden which others have voluntarily assumed and thus gains an advantage which others, who have restrained themselves, do not possess. This system, then, is one in which the rules establish a mutuality of benefit and burden and in which the benefits of noninterference are conditional upon the assumption of burdens.

Connecting punishment with the violation of these primary rules, and making public the provision for punishment, is both reasonable and just. First, it is only reasonable that those who voluntarily comply

with the rules be provided some assurance that they will not be assuming burdens which others are unprepared to assume. Their disposition to comply voluntarily will diminish as they learn that others are with impunity renouncing burdens they are assuming. Second, fairness dictates that a system in which benefits and burdens are equally distributed have a mechanism designed to prevent a maldistribution in the benefits and burdens. Thus, sanctions are attached to noncompliance with the primary rules so as to induce compliance with the primary rules among those who may be disinclined to obey. In this way the likelihood of an unfair distribution is diminished.

Third, it is just to punish those who have violated the rules and caused the unfair distribution of benefits and burdens. A person who violates the rules has something others have—the benefits of the system—but by renouncing what others have assumed, the burdens of self-restraint, he has acquired an unfair advantage. Matters are not even until this advantage is in some way erased. Another way of putting it is that he owes something to others, for he has something that does not rightfully belong to him. Justice—that is punishing such individuals—restores the equilibrium of benefits and burdens by taking from the individual what he owes, that is, exacting the debt. It is important to see that the equilibrium may be restored in another way. Forgiveness—with its legal analogue of a pardon—while not the righting of an unfair distribution by making one pay his debt is, nevertheless, a restoring of the equilibrium by forgiving the debt. Forgiveness may be viewed, at least in some types of cases, as a gift after the fact, erasing a debt, which had the gift been given before the fact, would not have created a debt. But the practice of pardoning has to proceed sensitively, for it may endanger in a way the practice of justice does not, the maintenance of an equilibrium of benefits and burdens. If all are indiscriminately pardoned less incentive is provided individuals to restrain their inclinations, thus increasing the incidence of persons taking what they do not deserve.

There are also in this system we are considering a variety of operative principles compliance with which provides some guarantee that the system of punishment does not itself promote an unfair distribution of benefits and burdens. For one thing, provision is made for a variety of defenses, each one of which can be said to have as its object diminishing the chances of forcibly depriving a person of benefits others have if that person has not derived an unfair advantage. A person has not derived an unfair advantage if he could not have restrained himself or if it is unreasonable to expect him to behave otherwise than he did. Sometimes the rules preclude punishment of classes of persons such as children. Sometimes they provide a defense if on a particular occasion a person lacked the capacity to conform his conduct to the rules. Thus, someone who in an epileptic seizure strikes another is excused. Punishment in these cases would be punishment of the innocent, punishment of those who do not voluntarily renounce a burden others have assumed. Punishment in such cases, then, would not equalize but rather cause an unfair distribution in benefits and burdens.

Along with principles providing defenses there are requirements that the rules be prospective and relatively clear so that persons have a fair opportunity to comply with the rules. There are, also, rules governing, among other matters, the burden of proof, who shall bear it and what it shall be, the prohibition on double jeopardy, and the privilege against self-incrimination. Justice requires conviction of the guilty, and requires their punishment, but in setting out to fulfill the demands of justice we may, of course, because we are not omniscient, cause injustice by convicting and punishing the innocent. The resolution arrived at in the system I am describing consists in weighing as the greater evil the punishment of the innocent. The primary function of the system of rules was to provide individuals with a sphere of interest immune from interference. Given this goal, it is determined to be a greater evil for society to interfere unjustifiably with an individual by depriving him of good than for the society to fail to punish those that have unjustifiably interfered.

Finally, because the primary rules are designed to benefit all and because the punishments prescribed for their violation are publicized and the defenses respected, there is some plausibility in the exaggerated claim that in choosing to do an act violative of the rules an individual has chosen to be punished. This way of putting matters brings to our attention the extent to which, when the system is as

I have described it, the criminal "has brought the punishment upon himself" in contrast to those cases where it would be misleading to say "he has brought it upon himself," cases, for example, where one does not know the rules or is punished in the absence of fault.

To summarize, then: first, there is a group of rules guiding the behavior of individuals in the community which establish spheres of interest immune from interference by others; second, provision is made for what is generally regarded as a deprivation of something of value if the rules are violated; third, the deprivations visited upon any person are justified by that person's having violated the rules; fourth, the deprivation, in this just system of punishment, is linked to rules that fairly distribute benefits and burdens and to procedures that strike some balance between not punishing the guilty and punishing the innocent, a class defined as those who have not voluntarily done acts violative of the law, in which it is evident that the evil of punishing the innocent is regarded as greater than the nonpunishment of the guilty.

At the core of many actual legal systems one finds, of course, rules and procedures of the kind I have sketched. It is obvious, though, that any ongoing legal system differs in significant respects from what I have presented here, containing "pockets of injustice."

I want now to sketch an extreme version of a set of institutions of a fundamentally different kind, institutions proceeding on a conception of man which appears to be basically at odds with that operative within a system of punishment.

Rules are promulgated in this system that prohibit certain types of injuries and harms.

In this world we are now to imagine when an individual harms another his conduct is to be regarded as a symptom of some pathological condition in the way a running nose is a symptom of a cold. Actions diverging from some conception of the normal are viewed as manifestations of a disease in the way in which we might today regard the arm and leg movements of an epileptic during a seizure. Actions conforming to what is normal are assimilated to the normal and healthy functioning of bodily organs. What a person does, then, is assimilated, on this conception, to what we believe today, or at least most of us believe today, a person undergoes. We draw a distinction between the operation of the kidney and raising an arm on request. This distinction between mere events or happenings and human actions is erased in our imagined system.[1] . . .

Let us elaborate on this assimilation of conduct of a certain kind to symptoms of a disease. First, there is something abnormal in both the case of conduct, such as killing another, and a symptom of a disease such as an irregular heart beat. Second, there are causes for this abnormality in action such that once we know of them we can explain the abnormality as we now can explain the symptoms of many physical diseases. The abnormality is looked upon as a happening with a causal explanation rather than an action for which there were reasons. Third, the causes that account for the abnormality interfere with the normal functioning of the body, or, in the case of killing with what is regarded as a normal functioning of an individual. . . .

With this view of man the institutions of social control respond, not with punishment, but with either preventive detention, in case of "carriers," or therapy in the case of those manifesting pathological symptoms. The logic of sickness implies the logic of therapy. . . . I am concerned now . . . with what the implications would be were the world indeed one of therapy . . . for I want to suggest tendencies of thought that arise when one is immersed in the ideology of disease and therapy.

First, punishment is the imposition upon a person who is believed to be at fault of something commonly believed to be a deprivation where that deprivation is justified by the person's guilty behavior. It is associated with resentment, for the guilty are those who have done what they had no right to do by failing to exercise restraint when they might have and where others have. Therapy is not a response to a person who is at fault. We respond to an individual, not because of what he has done, but because of some condition from which he is suffering. If he is no longer suffering from the condition, treatment no longer has a point. Punishment, then, focuses on the past; therapy on the present. Therapy is normally associated with compassion for what one undergoes, not resentment for what one has illegitimately done.

Second, with therapy, unlike punishment, we do not seek to deprive the person of something acknowledged as a good, but seek rather to help and to benefit the individual who is suffering by ministering to his illness in the hope that the person can be cured. The good we attempt to do is not a reward for desert. The individual suffering has not merited by his disease the good we seek to bestow upon him but has, because he is a creature that has the capacity to feel pain, a claim upon our sympathies and help.

Third, we saw with punishment that its justification was related to maintaining and restoring a fair distribution of benefits and burdens. Infliction of the prescribed punishment carries the implication, then, that one has "paid one's debt" to society, for the punishment is the taking from the person of something commonly recognized as valuable. It is this conception of "a debt owed" that may permit, as I suggested earlier, under certain conditions, the nonpunishment of the guilty, for operative within a system of punishment may be a concept analogous to forgiveness, namely pardoning. . . . What is clear is that the conceptions of "paying a debt" or "having a debt forgiven" or pardoning have no place in a system of therapy.

Fourth, with punishment there is an attempt at some equivalence between the advantage gained by the wrongdoer—partly based upon the seriousness of the interest invaded, partly on the state of mind with which the wrongful act was performed—and the punishment meted out. Thus, we can understand a prohibition on "cruel and unusual punishments" so that disproportionate pain and suffering are avoided. With therapy attempts at proportionality make no sense. It is perfectly plausible giving someone who kills a pill and treating for a lifetime within an institution one who has broken a dish and manifested accident proneness. We have the concept of "painful treatment." We do not have the concept of "cruel treatment." Because treatment is regarded as a benefit, though it may involve pain, it is natural that less restraint is exercised in bestowing it, than in inflicting punishment. Further, protests with respect to treatment are likely to be assimilated to the complaints of one whose leg must be amputated in order for him to live, and, thus, largely disregarded. To be sure, there is operative in the therapy world some conception of the "cure being worse than the disease," but if the disease is manifested in conduct harmful to others, and if being a normal operating human being is valued highly, there will naturally be considerable pressure to find the cure acceptable.

Fifth, the rules in our system of punishment governing conduct of individuals were rules violation of which involved either direct interference with others or the creation of a substantial risk of such interference. . . . Though we are interested in diminishing violations of the primary rules, we are not prepared to punish too many individuals who would never have violated the rules in order to achieve this aim. In a system motivated solely by a preventive and curative ideology there would be less reason to wait until symptoms manifest themselves in socially harmful conduct. It is understandable that we should wish at the earliest possible stage to arrest the development of the disease. In the punishment system, because we are dealing with deprivations, it is understandable that we should forbear from imposing them until we are quite sure of guilt. In the therapy system, dealing as it does with benefits, there is less reason for forbearance from treatment at an early stage.

Sixth, a variety of procedural safeguards we associate with punishment have less significance in a therapy system. To the degree objections to double jeopardy and self-incrimination are based on a wish to decrease the chances of the innocent being convicted and punished, a therapy system, unconcerned with this problem, would disregard such safeguards. When one is out to help people there is also little sense in urging that the burden of proof be on those providing the help. And there is less point to imposing the burden of proving that the conduct was pathological beyond a reasonable doubt. Further, a jury system which, within a system of justice, serves to make accommodations to the individual situation and to introduce a human element, would play no role or a minor one in a world where expertise is required in making determinations of disease and treatment.

In our system of punishment an attempt was made to maximize each individual's freedom of choice by first of all delimiting by rules certain

spheres of conduct immune from interference by others. The punishment associated with these primary rules paid deference to an individual's free choice by connecting punishment to a freely chosen act violative of the rules, thus giving some plausibility to the claim, as we saw, that what a person received by way of punishment he himself had chosen. With the world of disease and therapy all this changes and the individual's free choice ceases to be a determinative factor in how others respond to him. All those principles of our own legal system that minimize the chances of punishment of those who have not chosen to do acts violative of the rules tend to lose their point in the therapy system, for how we respond in a therapy system to a person is not conditioned upon what he has chosen but rather on what symptoms he has manifested or may manifest and what the best therapy for the disease is that is suggested by the symptoms.

Now, it is clear I think, that were we confronted with the alternatives I have sketched, between a system of just punishment and a thoroughgoing system of treatment, a system, that is, that did not reintroduce concepts appropriate to punishment, we could see the point in claiming that a person has a right to be punished, meaning by this that a person had a right to all those institutions and practices linked to punishment. For these would provide him with, among other things, a far greater ability to predict what would happen to him on the occurrence of certain events than the therapy system. There is the inestimable value to each of us of having the responses of others to us determined over a wide range of our lives by what we choose rather than what they choose. A person has a right to institutions that respect his choices. Our punishment system does; our therapy system does not.

Apart from those aspects of our therapy model which would relate to serious limitations on personal liberty, there are clearly objections of a more profound kind to the mode of thinking I have associated with the therapy model.

First, human beings pride themselves in having capacities that animals do not. A common way, for example, of arousing shame in a child is to compare the child's conduct to that of an animal. In a system where all actions are assimilated to happenings we are assimilated to creatures—indeed, it is more extreme than this—whom we have always thought possessed of less than we. Fundamental to our practice of praise and order of attainment is that one who can do more—one who is capable of more and one who does more is more worthy of respect and admiration. And we have thought of ourselves as capable where animals are not of making, of creating, among other things, ourselves. The conception of man I have outlined would provide us with a status that today, when our conduct is assimilated to it in moral criticism, we consider properly evocative of shame.

Second, if all human conduct is viewed as something men undergo, thrown into question would be the appropriateness of that extensive range of peculiarly human satisfactions that derive from a sense of achievement. For these satisfactions we shall have to substitute those mild satisfactions attendant upon a healthy well-functioning body. Contentment is our lot if we are fortunate; intense satisfaction at achievement is entirely inappropriate.

Third, in the therapy world nothing is earned and what we receive comes to us through compassion, or through a desire to control us. Resentment is out of place. We can take credit for nothing but must always regard ourselves—if there are selves left to regard once actions disappear—as fortunate recipients of benefits or unfortunate carriers of disease who must be controlled. We know that within our own world human beings who have been so regarded and who come to accept this view of themselves come to look upon themselves as worthless. When what we do is met with resentment, we are indirectly paid something of a compliment.

Fourth, attention should also be drawn to a peculiar evil that may be attendant upon regarding a man's actions as symptoms of disease. The logic of cure will push us toward forms of therapy that inevitably involve changes in the person made against his will. The evil in this would be most apparent in those cases where the agent, whose action is determined to be a manifestation of some disease, does not regard his action in this way. He believes that what he has done is, in fact, "right" but his conception of "normality" is not the therapeutically accepted one. When we treat an illness we normally treat a condition that the

person is not responsible for. He is "suffering" from some disease and we treat the condition, relieving the person of something preventing his normal functioning. When we begin treating persons for actions that have been chosen, we do not lift from the person something that is interfering with his normal functioning but we change the person so that he functions in a way regarded as normal by the current therapeutic community. We have to change him and his judgments of value. In doing this we display a lack of respect for the moral status of individuals, that is, a lack of respect for the reasoning and choices of individuals. They are but animals who must be conditioned. I think we can understand and, indeed, sympathize with a man's preferring death to being forcibly turned into what he is not.

Finally, perhaps most frightening of all would be the derogation in status of all protests to treatment. If someone believes that he has done something right, and if he protests being treated and changed, the protest will itself be regarded as a sign of some pathological condition, for who would not wish to be cured of an affliction? What this leads to are questions of an important kind about the effect of this conception of man upon what we now understand by reasoning. Here what a person takes to be a reasoned defense of an act is treated, as the action was, on the model of a happening of a pathological kind. Not just a person's acts are taken from him but also his attempt at a reasoned justification for the acts. In a system of punishment a person who has committed a crime may argue that what he did was right. We make him pay the price and we respect his right to retain the judgment he has made. A conception of pathology precludes this form of respect. . . .

I want also to make clear in concluding this section that I have argued, though very indirectly, not just for a right to a system of punishment, but for a right to be punished once there is in existence such a system. Thus, a man has the right to be punished rather than treated if he is guilty of some offense. And, indeed, one can imagine a case in which, even in the face of an offer of a pardon, a man claims and ought to have acknowledged his right to be punished.

NOTE

1. "When a man is suffering from an infectious disease, he is a danger to the community, and it is necessary to restrict his liberty of movement. But no one associates any idea of guilt with such a situation. On the contrary, he is an object of commiseration to his friends. Such steps as science recommends are taken to cure him of his disease, and he submits as a rule without reluctance to the curtailment of liberty involved meanwhile. The same method in spirit ought to be shown in the treatment of what is called 'crime.' "

   Bertrand Russell, *Roads to Freedom* (London: George Allen and Unwin Ltd., 1918), p. 135.

   "We do not hold people responsible for their reflexes—for example, for coughing in church. We hold them responsible for their operant behavior—for example, for whispering in church or remaining in church while coughing. But there are variables which are responsible for whispering as well as coughing, and these may be just as inexorable. When we recognize this, we are likely to drop the notion of responsibility altogether and with it the doctrine of free will as an inner causal agent."

   B. F. Skinner, *Science and Human Behavior* (1953), pp. 115–16.

   "Basically, criminality is but a symptom of insanity, using the term in its widest generic sense to express unacceptable social behavior based on unconscious motivation flowing from a disturbed instinctive and emotional life, whether this appears in frank psychoses, or in less obvious form in neuroses and unrecognized psychoses. . . . If criminals are products of early environmental influences in the same sense that psychotics and neurotics are, then it should be possible to reach them psychotherapeutically."

   Benjamin Karpman, "Criminal Psychodynamics," *Journal of Criminal Law and Criminology,* 47 (1956), p. 9.

   "We, the agents of society, must move to end the game of tit-for-tat and blow-for-blow in which the offender has foolishly and futilely engaged himself and us. We are not driven, as he is, to wild and impulsive actions. With knowledge comes power, and with power there is no need for the frightened vengeance of the old penology. In its place should go a quiet, dignified, therapeutic program for the rehabilitation of the disorganized one, if possible, the protection of society during the treatment period, and his guided return to useful citizenship, as soon as this can be effected."

   Karl Menninger, "Therapy, Not Punishment," *Harper's Magazine* (August 1959), pp. 63–64.

REVIEW AND DISCUSSION QUESTIONS

1. What are the basic features of a just system of punishment as Morris describes it? To what extent does our criminal justice system approximate Morris's idealized model?
2. Should Morris be considered a retributivist or a utilitarian?
3. What are the basic features of the disease-and-therapy model? What evidence do you see of the model in contemporary thinking about criminal justice? How representative are the authors quoted in Morris's footnote?
4. How do the two models differ in the way they view human beings? What difference is there between the two models in practice? Do you agree with Morris that, philosophically, there is something deeply objectionable about the therapy model? How might a proponent of that model respond to Morris?

The Moral Education Theory of Punishment

Jean Hampton

In this essay, Jean Hampton explains and argues for the moral education theory of punishment. Although this theory contains elements of both the retributivist and the utilitarian or deterrent approach to punishment, it is, she argues, a distinct theory that is superior to both and provides a complete justification of punishment in its own right. At the heart of the theory is the idea that punishment is a kind of moral communication because it teaches the wrongdoer a moral lesson. Because punishment educates criminals, it is good for them. Hampton's theory thus justifies punishment as a benefit to the wrongdoer, rather than as an evil that is justified for the greater good (utilitarianism) or that the criminal deserves to suffer (retributivism).

middle ground between deontological Engages with act & deals with best consequences.

I want to explore the promise of another theory of punishment which incorporates certain elements of the deterrence, retributivist, and rehabilitation views, but whose justification for punishment and whose formula for determining what punishment a wrongdoer deserves are distinctive and importantly different from the reasons and formulas characterizing the traditional rival theories. . . . Hence my discussion of the moral education theory in this paper is meant to develop it as a complete justification of punishment and to distinguish it from its traditional rivals. . . .

When does punishment by the state take place? The answer to this question seems simple: the state carries out punishment upon a person when he or she has broken a *law.* Yet the fact that the state's punishment always follows the transgression of a law is surely neither coincidental nor irrelevant to the understanding and justification of this practice. . . .

The threat, which specifies the infliction of pain if the imperative is not obeyed, gives people a nonmoral incentive, that is, the avoidance of pain, to refrain from the prohibited action. The state hopes this incentive will block a person's performance of the immoral action whenever the ethical incentive fails to do so. But insofar as the threat given in the law is designed to play this kind of "deterring" role, carrying out the threat, that is, punishing someone when he or she has broken the law, is, at least in part, a way of "making good" on the threat. The threat will only deter the disobedience of the state's orders if people believe there is a good chance the pain will be inflicted upon them after they commit the crime. But if the state punishes in order to make good on its threats, then the deterrence of future crime cannot be wholly irrelevant to the justification of punishment. . . .

But to say that the state's punishment is needed to prevent crime is not to commit oneself to the

deterrence justification of punishment—it all depends on what one takes prevention to entail. And, as Hegel says, if we aimed to prevent wrongdoing only by deterring its commission, we would be treating human beings in the same way that we treat dogs. Consider the kind of lesson an animal learns when, in an effort to leave a pasture, it runs up against an electrified fence. It experiences pain and is conditioned, after a series of encounters with the fence, to stay away from it and thus remain in the pasture. A human being in the same pasture will get the same message and learn the same lesson—"if you want to avoid pain, don't try to transgress the boundary marked by this fence." But, unlike the animal in the pasture, a human being will also be able to reflect on the reasons for that fence's being there, to theorize about *why* there is this barrier to his freedom.

Punishments are like electrified fences. At the very least they teach a person, via pain, that there is a "barrier" to the action she wants to do, and so, at the very least, they aim to deter. But because punishment "fences" are marking *moral* boundaries, the pain which these "fences" administer (or threaten to administer) conveys a larger message to beings who are able to reflect on the reasons for these barriers' existence: they convey that there is a barrier to these actions *because* they are morally wrong. Thus, according to the moral education theory, punishment is not intended as a way of conditioning a human being to do what society wants her to do (in the way that an animal is conditioned by an electrified fence to stay within a pasture); rather, the theory maintains that punishment is intended as a way of teaching the wrongdoer that the action she did (or wants to do) is forbidden because it is morally wrong and should not be done for that reason. The theory also regards that lesson as public, and thus as directed to the rest of society. When the state makes its criminal law and its enforcement practices known, it conveys an educative message not only to the convicted criminal but also to anyone else in the society who might be tempted to do what she did.

Comparing punishments to electrical fences helps to make clear how a certain kind of deterrent message is built into the larger moral point which punishment aims to convey. If one wants someone to understand that an offense is immoral, at the very least one has to convey to him or her that it is prohibited—that it ought not to occur. Pain is the way to convey that message. The pain says "Don't!" and gives the wrongdoer a reason for not performing the action again; an animal shocked by a fence gets the same kind of message and the same kind of incentive. But the state also wants to use the pain of punishment to get the human wrongdoer to reflect on the moral reasons for that barrier's existence, so that he will make the decision to reject the prohibited action for *moral* reasons, rather than for the self-interested reason of avoiding pain.

If those who are punished (or who watch the punishment take place) reject the moral message implicit in the punishment, at least they will learn from it that there is a barrier to the actions they committed (or are tempted to commit). Insofar as they choose to respond to their punishment (or the punishment of others) merely as a threat, it can keep them within moral boundaries in the same way that fences keep animals in a pasture. This deterrent effect of punishment is certainly welcome by the state whose role is to protect its citizens, and which has erected a "punishment barrier" to certain kinds of actions precisely because those actions will seriously harm its citizens. But on the moral education view, it is incorrect to regard simple deterrence as the aim of punishment; rather, to state it succinctly, the view maintains that punishment is justified as a way to prevent wrongdoing insofar as it can teach both wrongdoers and the public at large the moral reasons for *choosing* not to perform an offense. . . .

One distinction between the moral education view and the deterrence justification of punishment is that on the moral education view, the state is not concerned to use pain coercively so as to progressively eliminate certain types of behavior; rather, it is concerned to educate its citizens morally so that they choose not to engage in this behavior. Moreover, there is another important difference between the two views. On the deterrence view, the infliction of pain on certain individuals is justified as a way of promoting a larger social end. But critics of the deterrence view have pointed out that this is just to say that it is all right to *use* certain individuals to achieve a desirable social goal. The moral education theory, however, does not sanction the use of a criminal for social purposes; on the contrary, it attempts to justify punishment as a way to

benefit the person who will experience it, a way of helping him to gain moral knowledge if he chooses to listen. Of course other desirable social goals will be achieved through his punishment, goals which include the education of the larger community about the immorality of the offense, but none of these ends is to be achieved at the expense of the criminal. Instead the moral good which punishment attempts to accomplish within the wrongdoer makes it something which is done *for* him, not *to* him. . . .

Some readers might wonder how close the moral education view is to the old retribution theory. Indeed references in the literature to a view of this type frequently characterize it as a variant of retribution. Nonetheless, there are sharp and important differences between the two views. . . . Suffice to say now that whereas retributivism understands punishment as performing the rather metaphysical task of "negating the wrong" and "reasserting the right," the moral education theorist argues the there is a concrete moral *goal* which punishment should be designed to accomplish, and that goal includes the benefiting of the criminal himself. The state, as it punishes the lawbreaker, is trying to promote his moral personality; it realizes that "[h]is soul is in jeopardy as his victim's is not." Thus, it punishes him as a way of communicating a moral message to him, which he can accept or not, as he chooses. . . .

I would maintain that regarding punishment as a kind of moral communication is intuitively very natural and attractive. Consider, for example, what we say when we punish others: a father who punishes his child explains that he does so in order that the child "learn his lesson"; someone who has been physically harmed by another demands punishment "so that she will understand what she did to me"; a judge recently told a well-known user of cocaine that he was receiving a stiff sentence because his "matter-of-fact dabbling in cocaine . . . tells the whole world it is all right to use it." These kinds of remarks accompanying our punishment efforts suggest that our principal concern as we punish is to get the wrongdoer to stop doing the immoral action by communicating to her that her offense was immoral. And the last remark by the judge to the cocaine user shows that when the state punishes it is important that these communications be public, so that other members of society will hear the same moral message. Even people who seem to be seeking revenge on wrongdoers behave in ways which show that they too want to make a moral point not only to the wrongdoer, but to anyone else who will listen. The hero seeking revenge in a Western movie, for example, never simply shoots the bad guy in the back when he finds him—he always confronts the bad guy first (usually in the presence of other people) and tells him *why* he is about to die. Indeed, the movie would be unsatisfying if he didn't make that communication. And surely, the hero's desire to explain his actions is linked with his desire to convey to the bad guy and to others in society that the bad guy had "done him wrong."

Moreover, if one understands punishment as a moral message aimed at educating both the wrongdoer and the rest of society about the immorality of the offense, one has a powerful explanation (at least as powerful as the one offered by retributivism) of why victims so badly want their assailants punished. If the point of punishment is to convey to the criminal (and others) that the criminal *wronged* the victim, then punishment is implicitly recognizing the victim's plight, and honoring the moral claims of that individual. Punishment affirms as a *fact* that the victim has been wronged, and as a *fact* that he is owed a certain kind of treatment from others. Hence, on this view, it is natural for the victim to demand punishment because it is a way for the community to restore his moral status after it has been damaged by his assailant. . . .

Retributivists have a very interesting criticism of the moral education theory available to them. Granted, they might maintain, that punishment is connected with moral education, still this only provides an additional reason for punishing someone—it does not provide the fundamental justification of punishment. That fundamental justification, they would argue, is retributive: wrongdoers simply *deserve* to experience pain for the sake of the wrong they have committed. As Kant has argued, however much good one intends one's punishment to effect,

> yet it must first be justified in itself as punishment, i.e., as mere harm, so that if it stopped there, and the person punished could get no glimpse of kindness

> hidden behind this harshness, he must yet admit that justice was done him, and that his reward was perfectly suitable to his conduct.

Moreover, such modern retributivists as Walter Moberly have argued that it is only when the wrongdoer can assent to his punishment as already justified in virtue of his offense that the punishment can do him any good.

In a certain sense, Moberly's point is simply that a criminal will perceive his punishment as vindictive and vengeful unless he understands or accepts the fact that it is justified. But should the justification of punishment be cashed out in terms of the retributive concept of desert, given that it has been difficult for retributivists to say what they mean by the criminal's "deserving" punishment simply in virtue of his offense? Robert Nozick tries to cash out the retributive link between crime and "deserved" punishment by saying that the punishment represents a kind of "linkage" between the criminal and "right values." But why is inflicting pain on someone a way of effecting this linkage? Why isn't the infliction of a pleasurable experience for the sake of the crime just as good a way of linking the wrongdoer with these right values? And if Nozick explains the linkage of pain with crime by saying that the pain is necessary in order to communicate to the criminal that his action was wrong, he has answered the question but lost his retributive theory. Other philosophers, like Hegel, speak of punishment as a way of "annulling" or "canceling" the crime and hence "deserved" for that reason. But although Hegel's words have a nice metaphorical ring to them, it is hard to see how they can be given a literal force that will explain the retributivist concept of desert. As J. L. Mackie has written, insofar as punishment occurs after the crime, it certainly cannot cancel it—past events are not eliminated by later ones.

It is partly because retributivists have been at a loss to explain the notion of desert implicit in their theory of punishment that I have sought to propose and explore a completely nonretributivist justification of punishment. But my reasons for rejecting retributivism are deeper. The retributive position is that it is somehow morally appropriate to inflict pain for pain, to take an eye for an eye, a tooth for a tooth. But how is it ever morally appropriate to inflict one evil for the sake of another? How is the society that inflicts the second evil any different from the wrongdoer who has inflicted the first? He strikes first, they strike back; why is the second strike acceptable but the first not? Plato . . . insists that both harms are wrong; and Jesus attacks retributivism for similar reasons:

> You have learned that they were told, "Eye for eye, tooth for tooth." But what I tell you is this: Do not set yourself against the man who wrongs you. . . . You have heard that they were told "Love your neighbor, hate your enemy." But what I tell you is this: Love your enemy and pray for your persecutors; only so can you be children of your heavenly father, who makes the sun rise on good and bad alike, and sends the rain on the honest and dishonest.
>
> [Matt. 5:38–39, 43–46]

In other words, both reject retributivism because they insist that the only thing human beings "deserve" in this life is *good*, that no matter what evil a person has committed, no one is justified in doing further evil to her.

But if one accepts the idea that no one can ever deserve ill, can we hope to justify punishment? Yes, if punishment can be shown to be a good for the wrongdoer. The moral education theory makes just such an attempt to explain punishment as a good for those who experience it, as something done *for* them, not to them, something designed to achieve a goal that includes their own moral well-being. This is the justification of punishment the criminal needs to hear so that he can accept it as legitimate rather than dismiss it as vindictive. Therefore, my interest in the moral education theory is connected with my desire to justify punishment *as a good* for those who experience it, and to avoid any theoretical justification of punishment that would regard it as a deserved evil. Reflection on the punishment activities of those who truly love the people they punish, for example, the infliction of pain by a parent on a beloved but naughty child, suggests to me that punishment should not be justified as a deserved evil, but rather as an attempt, by someone who cares, to improve a wayward person. . . .

So the theory's point is this: Wrong occasions punishment not because pain deserves pain, but because evil deserves correction.

REVIEW AND DISCUSSION QUESTIONS

1. Although Hampton agrees that punishment is necessary to deter crime, she doesn't believe that this is what justifies punishment. Explain. What analogy does she draw between punishment and electrified fences?
2. Hampton contends that her theory, unlike the deterrence theory, does not inflict harm on criminals in order to benefit society. Are you persuaded by her reasoning? What about the idea that punishing the criminal sends a moral message to others?
3. How does Hampton's theory differ from retributivism? Assess her criticism of retributivism.
4. Punishment is almost always unpleasant. That's its point. But suppose the most effective way to morally educate wrongdoers did not involve doing anything unpleasant to them. Would this pose a challenge to Hampton's theory?
5. Do you agree that punishment should be for the good of the criminal? Has Hampton offered a viable alternative to the utilitarian and retributive justifications of punishment?

Restitution: A New Paradigm of Criminal Justice

Randy E. Barnett

Utilitarianism and retributivism usually occupy center stage in philosophical discussions of punishment. By contrast, Randy E. Barnett, professor of law at the Chicago-Kent College of Law, proposes thinking about criminal justice in terms of restitution. The idea of restitution focuses on the crime as an offense not against society, but against an individual. In a restitutional system, the criminal's debt is to the victim, not to society, and the victim's right to restitution determines the nature and extent of criminal liability.[1]

OUTLINE OF A NEW PARADIGM

The idea of restitution is actually quite simple. It views crime as an offense by one individual against the rights of another. The victim has suffered a loss. Justice consists of the culpable offender making good the loss he has caused. It calls for a complete refocusing of our image of crime. Kuhn would call it a "shift of worldview." Where we once saw an offense against society, we now see an offense against an individual victim. In a way, it is a common sense view of crime. *The armed robber did not rob society, he robbed the victim.* His debt, therefore, is not to society; it is to the victim. . . .

"Recompense or restitution is scarcely a punishment as long as it is merely a matter of returning stolen goods or money. . . . The point is not that the offender deserves to suffer; it is rather that the offended party desires compensation."[2] This represents the complete overthrow of the paradigm of punishment. No longer would the deterrence, reformation, disablement, or rehabilitation of the criminal be the guiding principle of the judicial system. The attainment of these goals would be incidental to, and as a result of, reparations paid to the victim. No longer would the criminal deliberately be made to suffer for his mistake. Making good that mistake is all that would be required. What follows is a possible scenario of such a system.

When a crime occurred and a suspect was apprehended, a trial court would attempt to determine his guilt or innocence. If found guilty, the criminal would be sentenced to make restitution to the victim. If a criminal is able to make restitution immediately, he may do so. This would discharge his liability. If he were unable to make restitution, but were found by the court to be trustworthy, he would be permitted to remain at his job (or find a new one) while paying

restitution out of his future wages. This would entail a legal claim against future wages. Failure to pay could result in garnishment or a new type of confinement.

If it is found that the criminal is not trustworthy, or that he is unable to gain employment, he would be confined to an employment project. This would be an industrial enterprise, preferably run by a private concern, which would produce actual goods or services. The level of security at each employment project would vary according to the behavior of the offenders. Since the costs would be lower, inmates at a lower-security project would receive higher wages. There is no reason why many workers could not be permitted to live with their families inside or outside the facility, depending, again, on the trustworthiness of the offender. Room and board would be deducted from the wages first, then a certain amount for restitution. Anything over that amount the worker could keep or apply toward further restitution, thus hastening his release. If a worker refused to work, he would be unable to pay for his maintenance, and therefore would not in principle be entitled to it. If he did not make restitution he could not be released. The exact arrangement which would best provide for high productivity, minimal security, and maximum incentive to work and repay the victim cannot be determined in advance. Experience is bound to yield some plans superior to others. In fact, the experimentation has already begun.

While this might be the basic system, all sorts of refinements are conceivable, and certainly many more will be invented as needs arise. A few examples might be illuminating. With such a system of repayment, victim *crime insurance* would be more economically feasible than at present and highly desirable. The cost of awards would be offset by the insurance company's right to restitution in place of the victim (right of subrogation). The insurance company would be better suited to supervise the offender and mark his progress than would the victim. To obtain an earlier recovery, it could be expected to innovate so as to enable the worker to repay more quickly (and, as a result, be released that much sooner). The insurance companies might even underwrite the employment projects themselves as well as related industries which would employ the skilled worker after his release. Any successful effort on their part to reduce crime and recidivism would result in fewer claims and lower premiums. The benefit of this insurance scheme for the victim is immediate compensation, conditional on the victim's continued cooperation with the authorities for the arrest and conviction of the suspect. In addition, the centralization of victim claims would, arguably, lead to efficiencies which would permit the pooling of small claims against a common offender.

Another highly useful refinement would be *direct arbitration* between victim and criminal. This would serve as a sort of healthy substitute for plea bargaining. By allowing the guilty criminal to negotiate a reduced payment in return for a guilty plea, the victim (or his insurance company) would be saved the risk of an adverse finding at trial and any possible additional expense that might result. This would also allow an indigent criminal to substitute personal services for monetary payments if all parties agreed. . . .

Something analogous to the medieval Irish system of *sureties* might be employed as well.[3] Such a system would allow a concerned person, group, or company to make restitution (provided the offender agrees to this). The worker might then be released in the custody of the surety. If the surety had made restitution, the offender would owe restitution to the surety who might enforce the whole claim or show mercy. Of course, the more violent and unreliable the offender, the more serious and costly the offense, the less likely it would be that anyone would take the risk. But for first offenders, good workers, or others that charitable interests found deserving (or perhaps unjustly convicted) this would provide an avenue of respite.

RESTITUTION AND RIGHTS

These three possible refinements clearly illustrate the flexibility of a restitutional system. It may be less apparent that this flexibility is *inherent* to the restitutional paradigm. Restitution recognizes rights in the victim, and this is a principal source of its strength. The nature and limit of the victim's right to restitution at the same time defines the nature and limit of the criminal liability. In this way, the aggressive action of the criminal creates a *debt* to the victim. The recognition of rights and obligations make possible many

innovative arrangements. Subrogation, arbitration, and suretyship are three examples mentioned above. They are possible because this right to compensation[4] is considered the property of the victim and can therefore be delegated, assigned, inherited, or bestowed. One could determine in advance who would acquire the right to any restitution which he himself might be unable to collect.

The natural owner of an unenforced death claim would be an insurance company that had insured the deceased. The suggestion has been made that a person might thus increase his personal safety by insuring with a company well known for tracking down those who injure its policy holders. In fact, the partial purpose of some insurance schemes might be to provide the funds with which to track down the malefactor. The insurance company, having paid the beneficiaries would "stand in their shoes." It would remain possible, of course, to simply assign or devise the right directly to the beneficiaries, but this would put the burden of enforcement on persons likely to be unsuited to the task.

If one accepts the Lockean trichotomy of property ownership,[5] that is, acquiring property via exchange, gifts, and *homesteading* (mixing one's labor with previously unowned land or objects), the possibility arises that upon a person's wrongful death, in the absence of any heirs or assignees, his right to compensation becomes unowned property. The right could then be claimed (homesteaded) by anyone willing to go to the trouble of catching and prosecuting the criminal. Firms might specialize in this sort of activity, or large insurance companies might make the effort as a kind of "loss leader" for public relations purposes.

This does, however, lead to a potentially serious problem with the restitutional paradigm: what exactly constitutes "restitution"? What is the *standard* by which compensation is to be made? . . . Any such problem facing the restitutional paradigm faces civil damage suits as well. The method by which this problem is dealt with in civil cases could be applied to restitution cases. But while this is certainly true, it may be that this problem has not been adequately handled in civil damage suits either.

Restitution in cases of crimes against property is a manageable problem. Modern contract and tort doctrines of restitution are adequate. The difficulty lies in cases of personal injury or death. How can you put a price on life or limb, pain or suffering? Is not any attempt to do so of necessity arbitrary? It must be admitted that a fully satisfactory solution to this problem is lacking, but it should also be stressed that this dilemma, though serious, has little impact on the bulk of our case in favor of a restitutional paradigm. It is possible that no paradigm of criminal justice can solve every problem, yet the restitutional approach remains far superior to the paradigm of punishment or any other conceivable rival.

This difficulty arises because certain property is unique and irreplaceable. As a result, it is impossible to approximate a "market" or "exchange" value expressed in monetary terms. Just as there is no rational relationship between a wrongfully taken life and ten years in prison, there is little relationship between that same life and $20,000. Still, the nature of this possibly insoluble puzzle reveals a restitutional approach theoretically superior to punishment. For it must be acknowledged that a real, tangible loss *has* occurred. The problem is only one of incommensurability. Restitution provides *some* tangible, albeit inadequate, compensation for personal injury. Punishment provides none at all.[6] . . .

ADVANTAGES OF A RESTITUTIONAL SYSTEM

1. The first and most obvious advantage is the assistance provided to victims of crime. They may have suffered an emotional, physical, or financial loss. Restitution would not change the fact that a possibly traumatic crime has occurred (just as the award of damages does not undo tortious conduct). Restitution, however, would make the resulting loss easier to bear for both victims and their families. At the same time, restitution would avoid a major pitfall of victim compensation/welfare plans: Since it is the criminal who must pay, the possibility of collusion between victim and criminal to collect "damages" from the state would be all but eliminated.

2. The possibility of receiving compensation would encourage victims to report crimes and to appear at trial. This is particularly true if there were a crime insurance scheme which contractually committed the policyholder to testify as a condition for payment, thus rendering unnecessary

oppressive and potentially tyrannical subpoenas and contempt citations. Even the actual reporting of the crime to police is likely to be a prerequisite for compensation. Such a requirement in auto theft insurance policies has made car thefts the most fully reported crime in the United States. Furthermore, insurance companies which paid the claim would have a strong incentive to see that the criminal was apprehended and convicted. Their pressure and assistance would make the proper functioning of law enforcement officials all the more likely.

3. Psychologist Albert Eglash has long argued that restitution would aid in the rehabilitation of criminals: "Restitution is something an inmate does, not something done for or to him. . . . Being reparative, restitution can alleviate guilt and anxiety, which can otherwise precipitate further offenses."[7] Restitution, says Eglash, is an active effortful role on the part of the offender. It is socially constructive, thereby contributing to the offender's self-esteem. It is related to the offense and may thereby redirect the thoughts which motivated the offense. It is reparative, restorative, and may actually leave the situation better than it was before the crime, both for the criminal and victim.

4. This is a genuinely "self-determinative" sentence. The worker would know that the length of his confinement was in his own hands. The harder he worked, the faster he would make restitution. He would be the master of his fate and would have to face that responsibility. This would encourage useful, productive activity and instill a conception of reward for good behavior and hard work. Compare this with the current probationary system and "indeterminate sentencing" where the decision for release is made by the prison bureaucracy, based only (if fairly administered) on "good behavior"; that is, passive acquiescence to prison discipline. Also, the fact that the worker would be acquiring *marketable* skills rather than more skillful methods of crime should help to reduce the shocking rate of recidivism.

5. The savings to taxpayers would be enormous. No longer would the innocent taxpayer pay for the apprehension and internment of the guilty. The cost of arrest, trial, and internment would be borne by the criminal himself. In addition, since now-idle inmates would become productive workers (able, perhaps, to support their families), the entire economy would benefit from the increase in overall production.

6. Crime would no longer pay. Criminals, particularly shrewd white-collar criminals, would know that they could not dispose of the proceeds of their crime and, if caught, simply serve time. They would have to make full restitution plus enforcement and legal costs, thereby greatly increasing the incentive to prosecute. While this would not eliminate such crime it would make it rougher on certain types of criminals, like bank and corporation officials, who harm many by their acts with a virtual assurance of lenient legal sanctions. It might also encourage such criminals to keep the money around for a while so that, if caught, they could repay more easily. This would make a full recovery more likely.

A restitutional system of justice would benefit the victim, the criminal, and the taxpayer. The humanitarian goals of proportionate punishment, rehabilitation, and victim compensation are dealt with on a *fundamental* level making their achievement more likely. In short, the paradigm of restitution would benefit all but the entrenched penal bureaucracy and enhance justice at the same time. What then is there to stop us from overthrowing the paradigm of punishment and its penal system and putting in its place this more efficient, more humane, and more just system? The proponents of punishment and others have a few powerful counterarguments. It is to these we now turn.

OBJECTIONS TO RESTITUTION

1. ***Practical criticisms of restitution.*** It might be objected that "crimes disturb and offend not only those who are directly their victim, but also the whole social order."[8] Because of this, society, that is, individuals other than the victim, deserves some satisfaction from the offender. Restitution, it is argued, will not satisfy the lust for revenge felt by the victim or the "community's sense of justice." This criticism appears to be overdrawn. Today most members of the community are mere

spectators of the criminal justice system, and this is largely true even of the victim. One major reform being urged presently is more victim involvement in the criminal justice process. The restitution proposal would necessitate this involvement. And while the public generally takes the view that officials should be tougher on criminals, with "tougher" taken by nearly everyone to mean more severe in punishing, one must view this "social fact" in light of the lack of a known alternative. The real test of public sympathies would be to see which sanction people would choose: incarceration of the criminal for a given number of years or the criminal's being compelled to make restitution to the victim. While the public's choice is not clearly predictable, neither can it be assumed that it would reject restitution. There is some evidence to the contrary.

This brings us to a second practical objection: that monetary sanctions are insufficient deterrents to crime. Again, this is something to be discovered, not something to be assumed. There are a number of reasons to believe that our *current* system of punishment does not adequately deter. . . . In fact, many have argued that the deterrent value of sanctions has less to do with *severity* than with *certainty*, and the preceding considerations indicate that law enforcement would be more certain under a restitutional system. In the final analysis, however, it is irrelevant to argue that more crimes may be committed if our proposal leaves the victim better off. It must be remembered: *Our goal is not the suppression of crime; it is doing justice to victims.*

A practical consideration which merits considerable future attention is the feasibility of the employment project proposal. A number of questions can be raised. At first blush, it seems naively optimistic to suppose that offenders will be able or willing to work at all, much less earn their keep and pay reparations as well. On the contrary, this argument continues, individuals turn to crime precisely because they lack the skills which the restitutional plan assumes they have. Even if these workers have the skills, but refuse to work, what could be done? Would not the use of force to compel compliance be tantamount to slavery?

. . . [O]ne can advance several responses. First, the problem as usually posed assumes the offender to be highly irrational and possibly mentally unbalanced. There is no denying that some segment of the criminal population fits the former description. What this approach neglects, however, is the possibility that many criminals are making rational choices within an irrational and unjust political system. Specifically I refer to the myriad laws and regulations which make it difficult for the unskilled or persons of transitory outlook to find legal employment. I refer also to the laws which deny legality to the types of services which are in particular demand in economically impoverished communities. Is it "irrational" to choose to steal or rob when one is virtually foreclosed from the legal opportunity to do otherwise? Another possibility is that the criminal chooses crime not because of foreclosure, but because he enjoys and obtains satisfaction from a criminal way of life. Though morally repugnant, this is hardly irrational.

Furthermore, it no longer can be denied that contact with the current criminal justice system is itself especially damaging among juveniles. The offenders who are hopelessly committed to criminal behavior are not usually the newcomers to crime but those who have had repeated exposure to the penal system. In Kuhn's words, "Existing institutions have ceased to meet the problems posed by an environment *they have in part created.*" While a restitutionary system might not change these hard-core offenders, it could, by the early implementation of sanctions perceived by the criminal to be just, break the vicious circle which in large part accounts for their existence.

Finally, if offenders could not or would not make restitution, then the logical and just result of their refusal would be confinement until they could or would. Such an outcome would be entirely in their hands. While this "solution" does not suggest who should justly pay for this confinement, the problem is not unique to a restitutionary system. In this and other areas of possible difficulty we must seek guidance from existing pilot programs as well as from the burgeoning research in this area and in victimology in general.

2. *Distributionary criticisms of restitution.* There remains one criticism of restitution which is the most obvious and the most difficult with which to deal. Simply stated, it takes the following form: "Doesn't this mean that rich people will be able to commit crimes with impunity if they can afford it? Isn't this unfair?" The *practical* aspect of this objection is that whatever deterrent effect restitution payments may have, they will be less for those most able to pay. The *moral* aspect is that whatever retributive or penal effect restitution payments may have they will be less for those who are well off. Some concept of equality of justice underlies both considerations.

Critics of restitution fail to realize that the "cost" of crime will be quite high. In addition to compensation for pain and suffering, the criminal must pay for the cost of his apprehension, the cost of the trial, and the legal expenditures of *both* sides. This should make even an unscrupulous wealthy person think twice about committing a crime. The response to this is that we cannot have it both ways. If the fines would be high enough to bother the rich, then they would be so high that a project worker would have no chance of earning that much and would, therefore, have no incentive to work at all. If, on the other hand, you lower the price of crime by ignoring all its costs, you fail to deter the rich or fully compensate the victim.

This is where the option of arbitration and victim crime insurance becomes of practical importance. If the victim is uninsured, he is unlikely to recover for all costs of a very severe crime from a poor, unskilled criminal, since even in an employment project the criminal might be unable to earn enough. If he had no hope of earning his release, he would have little incentive to work very hard beyond paying for his own maintenance. The victim would end up with less than if he had "settled" the case for the lesser amount which a project worker could reasonably be expected to earn. If, however, the victim had full-coverage criminal insurance, he would recover his damages in full, and the insurance company would absorb any disparity between full compensation and maximal employment project worker's output. This cost would be reflected in premium prices, enabling the insurance company which settled cases at an amount which increased the recovery from the criminal to offer the lowest rates. Eventually a "maximum" feasible fine for project workers would be determined based on these considerations. The "rich," on the other hand, would naturally have to pay in full. This arrangement would solve the practical problem, but it should not be thought of as an imperative of the restitutional paradigm.

. . . The paradigm of restitution calls not for the (equal) hurting of criminals, but for restitution to victims. Any appeal to "inadequate suffering" is a reversion to the paradigm of punishment, and by varying the sanction for crimes of the same magnitude according to the economic status of the offender it reveals its own inequity. *Equality of justice means equal treatment of victims.* It should not matter to the victim if his attacker was rich or poor. His plight is the same regardless. Any reduction of criminal liability because of reduced earning power would be for practical, not moral, reasons.

Equality of justice derives from the fact that the rights of men should be equally enforced and respected. Restitution recognizes a victim's right to compensation for damages from the party responsible. Equality of justice, therefore, calls for equal enforcement of each victim's right to restitution. *Even if necessary or expedient, any lessening of payment to the victim because of the qualities of the criminal is a violation of that victim's rights and an inequality of justice.* Any such expedient settlement is only a recognition that an imperfect world may make possible only imperfect justice. As a practical matter, a restitutional standard gives victims an enormous incentive to pursue wealthy criminals since they can afford quick, full compensation. Contrast this with the present system where the preference given the wealthy is so prevalent that most victims simply assume that nothing will be done.

The paradigm of restitution, to reiterate, is neither a panacea for crime nor a blueprint for utopia. Panaceas and utopias are not for humankind. We must live in a less than perfect world with less than perfect people. Restitution opens the possibility of an improved and more just society. The old paradigm of punishment, even reformed, simply cannot offer this promise.

NOTES

1. For further discussion of this proposal, *see* Randy E. Barnett, "The Justice of Restitution," *25 American Journal of Jurisprudence* 117 (1980); Randy E. Barnett, "Pursuing Justice in a Free Society: Part One—Power vs. Liberty," *Criminal Justice Ethics*, Summer/Fall 1985, at 50; Randy E. Barnett, "Pursuing Justice in a Free Society: Part Two—Crime Prevention and the Legal Order," *Criminal Justice Ethics*, Winter/Spring 1986, at 30.
2. Kaufmann, p. 55.
3. For a description of the Irish system, see Joseph R. Peden, "Property Rights in Medieval Ireland: Celtic Law versus Church and State" (paper presented at the Symposium on the Origins and Development of Property Rights, University of San Francisco, January 1973); for a theoretical discussion of a similar proposal, see Spencer, pp. 182–86.
4. Or, perhaps more accurately, the compensation itself.
5. For a brief explanation of this concept and several of its possible applications, see Murray N. Rothbard, "Justice and Property Rights" in *Property in a Humane Economy*, ed. Samuel L. Blumenfeld (La Salle, Ill.: Open Court Publishing Co., 1974), pp. 101–22.
6. That the "spiritual" satisfaction which punishment may or may not provide is to be recognized as a legitimate form of "compensation" is a claim retributionists must defend.
7. Albert Eglash, "Creative Restitution: Some Suggestions for Prison Rehabilitation Programs," *American Journal of Correction* 40 (November–December 1958): 20.
8. Del Vecchio, p. 198.

REVIEW AND DISCUSSION QUESTIONS

1. What is the difference between restitution and punishment?
2. With reference to Barnett's discussion of insurance, arbitration, and sureties, what sort of a system of restitution would be most feasible? In practice, how could we determine the exact compensation owed to a victim?
3. How, in Barnett's view, does restitution benefit criminals?
4. What are the advantages to taxpayers of Barnett's system?
5. What do you see as the most important objections to a restitutional system? How would Brandt, Morris, and Hampton respond to Barnett's proposal? Whose approach to punishment and criminal justice do you find the most attractive?

CHAPTER 13

The Rights of Defendants

The vast majority of criminal cases are settled by plea bargains. Because trials are expensive and time consuming, without plea bargaining our criminal justice system would probably grind to a halt. Kenneth Kipnis argues, however, that plea bargains may violate one's right not to incriminate oneself and that they rarely produce a just outcome.

Among our most important legal rights as citizens are those that protect us from arbitrary police conduct and from unjust treatment by the criminal justice system. Many people, however, view skeptically the rights of criminal defendants because they believe that those whom the police arrest are usually guilty as charged. In "Convicting the Innocent," James McCloskey challenges this belief, arguing that at least 10 percent of those convicted of serious crimes are completely innocent.

On the other hand, most people believe that the police are sometimes justified in employing morally or legally questionable means to solve or prevent crimes. Carl Klockars explores this dilemma, which he calls the Dirty Harry Problem.

The Fourth Amendment protects us from unreasonable searches and seizures, requiring that the police have probable cause before a warrant is issued. To enforce this right, the Supreme Court has held that illegally obtained evidence should be excluded from criminal trials. This is called the exclusionary rule, and in their essays, Malcolm Richard Wilkey and Stephen H. Sachs debate its wisdom. The Fifth Amendment provides that no one "shall be compelled in any criminal case to be a witness against himself," and the Sixth Amendment affords us the right to counsel, but specifying the exact nature and limits of these rights in real life situations is not always easy. In the controversial *Miranda, Williams,* and *Innis* cases, the Supreme Court has tried to spell out the exact procedural safeguards that the police must follow to respect these constitutional protections. In the later *Tobias* case, a federal court probed the question of when the actions of law enforcement officials constitute entrapment and thus violate a criminal defendant's rights.

Criminal Justice and the Negotiated Plea

Kenneth Kipnis

Most criminal cases are settled by negotiated pleas ("plea bargains"). In fact, our criminal justice system today depends on most defendants agreeing not to go to trial. Kenneth Kipnis, professor of philosophy at the University of Hawaii—Manoa, criticizes the institution of plea bargaining on the ground that it involves something comparable to duress and thus erodes our right to be free from compelled self-incrimination. Furthermore, innocent people may be forced to "cop a plea," and because even guilty individuals are not given the punishment they deserve, plea bargaining rarely, if ever, produces a just outcome.

In recent years it has become apparent to many that, in practice, the criminal justice system in the United States does not operate as we thought it did. The conviction secured through jury trial, so familiar in countless novels, films, and television programs, is beginning to be seen as the aberration it has become. What has replaced the jury's verdict is the negotiated plea. In these "plea bargains" the defendant agrees to plead guilty in exchange for discretionary consideration on the part of the state. Generally, this consideration amounts to some kind of assurance of a minimal sentence. . . . It is at present a commonplace that plea bargaining could not be eliminated without substantial alterations in our criminal justice system.

Plea bargaining involves negotiations between the defendant (through an attorney in the standard case) and the prosecutor as to the conditions under which the defendant will enter a guilty plea. Both sides have bargaining power in these negotiations. The prosecutor is ordinarily burdened with cases and does not have the wherewithal to bring more than a fraction of them to trial. Often there is not sufficient evidence to ensure a jury's conviction. Most important, the prosecutor is typically under administrative and political pressure to dispose of cases and to secure convictions as efficiently as possible. If the defendant exercises the constitutional right to a jury trial, the prosecutor must decide whether to drop the charges entirely or to expend scarce resources to bring the case to trial. Since neither prospect is attractive, prosecutors typically exercise their broad discretion to induce defendants to waive trial and to plead guilty.

From the defendant's point of view, such prosecutorial discretion has two aspects: it darkens the prospect of going to trial as it brightens the prospect of pleading guilty. Before negotiating, a prosecutor may improve his bargaining position by "overcharging" defendants or by developing a reputation for severity in the sentences he recommends to judges. Such steps greatly increase the punishment that the defendant must expect if convicted at trial. On the other hand, the state may offer to reduce or to drop some charges, or to recommend leniency to the judge if the defendant agrees to plead guilty. These steps minimize the punishment that will result from a guilty plea. Though the exercise of prosecutorial discretion to secure pleas of guilty may differ somewhat in certain jurisdictions and in particular cases, the broad outlines are as described.

Of course a defendant can always reject any offer of concessions and challenge the state to prove its case. A skilled defense attorney can do much to force the prosecutor to expend resources in bringing a case to trial. But the trial route is rarely taken by defendants. Apart from prosecutorial pressure, other factors may contribute to a defendant's willingness to plead guilty: feelings of guilt which may or may not be connected with the charged crime; the discomforts of the pretrial lockup as against the comparatively better facilities of a penitentiary; the costs of going to trial as against the often cheaper option of consenting to a plea; a willingness or unwillingness to lie; and the delays which are almost always present in awaiting trial, delays which the defendant may sit out in jail in a kind of preconviction imprisonment which may not be credited to a postconviction sentence. It is not surprising that the right to a trial by jury is rarely exercised. . . .

No deliberative body ever decided that we would have a system in which the disposition of criminal cases is typically the result of negotiations between the prosecutor and the defendant's attorney on the conditions under which the defendant would waive trial and plead guilty to a mutually acceptable charge. No legislature ever voted to adopt a procedure in which defendants who are convicted after trial typically receive sentences far greater than those received by defendants charged with similar offenses but pleading guilty. The practice of plea bargaining has evolved in the unregulated interstices of our criminal justice system. Its development has not gone unnoticed. There is now a substantial literature on the legality and propriety of plea bargaining. But though philosophers do not often treat issues arising in the area of criminal procedure, there are problems here that cry for our attention. In the preceding pages I have been concerned to sketch the institution of plea bargaining. In what follows I will raise some serious questions about it that should concern us. I will first discuss generally the intrinsic fairness of plea bargains and then, in the final section, I will examine critically the place of such bargains in the criminal justice system.

1

As one goes through the literature on plea bargaining one gets the impression that market forces are at work in this unlikely context. The terms "bargain" and "negotiation" suggest this. One can see the law of supply and demand operating in that, other things being equal, if there are too many defendants who want to go to trial, prosecutors will have to concede more in order to get the guilty pleas that they need to clear their case load. And if the number of prosecutors and courts goes up, prosecutors will be able to concede less. Against this background it is not surprising to find one commentator noting: "In some places a 'going rate' is established under which a given charge will automatically be broken down to a given lesser offense with the recommendation of a given lesser sentence." Prosecutors, like retailers before them, have begun to appreciate the efficiency of the fixed-price approach.

The plea bargain in the economy of criminal justice has many of the important features of the contract in commercial transactions. In both institutions offers are made and accepted, entitlements are given up and obtained, and the notion of an exchange, ideally a fair one, is present to both parties. Indeed one detects something of the color of consumer protection law in a few of the decisions on plea bargaining. In *Bailey v. MacDougal* the court held that "a guilty plea cannot be accepted unless the defendant understands its consequences." And in *Santo Bello v. New York* the court secured a defendant's entitlement to a prosecutorial concession when a second prosecutor replaced the one who had made the promise. . . . Though plea bargains may not be seen as contracts by the parties, agreements like them are the stuff of contract case law. While I will not argue that plea bargains are contracts (or even that they should be treated as such), I do think it proper to look to contract law for help in evaluating the justice of such agreements.

The law of contracts serves to give legal effect to certain bargain-promises. In particular, it specifies conditions that must be satisfied by bargain-promises before the law will recognize and enforce them as contracts. As an example, we could look at that part of the law of contracts which treats duress. Where one party wrongfully compels another to consent to the terms of an agreement the resulting bargain has no legal effect. Dan B. Dobbs, a commentator on the law in this area, describes the elements of duress as follows: "The defendant's act must be wrongful in some attenuated sense; it must operate coercively upon the will of the plaintiff, judged subjectively, and the plaintiff must have no adequate remedy to avoid the coercion except to give in. . . . The earlier requirement that the coercion must have been the kind that would coerce a reasonable man, or even a brave one, is now generally dispensed with, and it is enough if it in fact coerced a spineless plaintiff." Coercion is not the same as fraud, nor is it confined to cases in which a defendant is physically compelled to assent. In Dobbs' words: "The victim of duress knows the facts but is forced by hard choices to act against his will." The paradigm case of duress is the agreement made at gunpoint. Facing a mortal threat, one readily agrees to hand over the cash. But despite such consent, the rules of duress work to void the effects of such agreements. There is no legal

obligation to hand over the cash and, having given it over, entitlement to the money is not lost. The gunman has no legal right to retain possession even if he adheres to his end of the bargain and scraps his murderous plans.

Judges have long been required to see to it that guilty pleas are entered voluntarily. And one would expect that, if duress is present in the plea-bargaining situation, then, just as the handing over of cash to the gunman is void of legal effect (as far as entitlement to the money is concerned), so no legal consequences should flow from the plea of guilty which is the product of duress. However, Rule 11 of the Federal Rules of Criminal Procedure requires the court to insure that a plea of guilty (or nolo contendere) is voluntary by "addressing the defendant personally in open court, determining that the plea is voluntary and not the result of force or promises *apart from a plea agreement*" (emphasis added). In two important cases (*North Carolina v. Alford* and *Brady v. United States*) defendants agreed to plead guilty in order to avoid probable death sentences. Both accepted very long prison sentences. In both cases the Supreme Court decided that guilty pleas so entered were voluntary (though Brennan, Douglas, and Marshall dissented). In his dissent in *Alford,* Brennan writes: ". . . the facts set out in the majority opinion demonstrate that Alford was 'so gripped by fear of the death penalty' that his decision to plead guilty was not voluntary but was that product of duress as much so as choice reflecting physical constraint." In footnote 2 of the *Alford* opinion, the Court sets out the defendant's testimony given at the time of the entry of his plea of guilty before the trial court. That testimony deserves examination: "I pleaded guilty on second degree murder because they said there is too much evidence, but I ain't shot no man, but I take the fault for the other man. We never had an argument in our life and I just pleaded guilty because they said if I didn't they would gas me for it, and that is all." The rule to be followed in such cases is set out in *Brady:* "A plea of guilty entered by one fully aware of the direct consequences, including the actual value of any commitments made to him by the court, prosecutor or his own counsel, must stand unless induced by threats (or promises to discontinue improper harassment), misrepresentation (including unfilled or unfillable promises), or perhaps by promises that are by their very nature improper as having no proper relationship to the prosecutor's business (e.g. bribes)." Case law and the Federal Rules both hold that the standard exercise of prosecutorial discretion in order to secure a plea of guilty cannot be used to prove that such a plea is [in]voluntary. Even where the defendant enters a guilty plea in order to avert his death at the hands of the state, as in *Alford,* the Court has not seen involuntariness. Nevertheless, it may be true that some guilty pleas are involuntary in virtue of prosecutorial inducement considered proper by the Supreme Court.

Regarding the elements of duress, let us compare the gunman situation with an example of plea bargaining in order to examine the voluntariness of the latter. Albert W. Alschuler, author of one of the most thorough studies of plea bargaining, describes an actual case:

> San Francisco defense attorney Benjamin M. Davis recently represented a man charged with kidnapping and forcible rape. The defendant was innocent, Davis says, and after investigating the case Davis was confident of an acquittal. The prosecutor, who seems to have shared the defense attorney's opinion on this point, offered to permit a guilty plea to simple battery. Conviction on this charge would not have led to a greater sentence than thirty days' imprisonment, and there was every likelihood that the defendant would be granted probation. When Davis informed his client of this offer, he emphasized that conviction at trial seemed highly improbable.
>
> The defendant's reply was simple: "I can't take the chance."

Both the gunman and the prosecutor require persons to make hard choices between a very certain smaller imposition and an uncertain greater imposition. In the gunman situation I must choose between the very certain loss of money and the difficult-to-assess probability that my assailant is willing and able to kill me if I resist. As a defendant I am forced to choose between a very certain smaller punishment and a substantially greater punishment with a difficult-to-assess probability. As the size of the certain smaller imposition comes down and as the magnitude and probability of the larger imposition increases, it becomes more and more reasonable to choose the former. This is what seems to be occurring in Alschuler's example:

"Davis reports that he is uncomfortable when he permits innocent defendants to plead guilty; but in this case it would have been playing God to stand in the defendant's way. The attorney's assessment of the outcome at trial can always be wrong, and it is hard to tell a defendant that 'professional ethics' require a course that may ruin his life." Davis's client must decide whether to accept a very certain, very minor punishment or to chance a ruined life. Of course the gunman's victim can try to overpower his assailant and the defendant can attempt to clear himself at trial. But the same considerations that will drive reasonable people to give in to the gunman compel one to accept the prosecutor's offer. Applying the second and third elements of duress, one can see that, like the gunman's acts, the acts of the prosecutor can "operate coercively upon the will of the plaintiff, judged subjectively," and both the gunman's victim and the defendant may "have no adequate remedy to avoid the coercion except to give in." In both cases reasonable persons might well conclude (after considering the gunman's lethal weapon or the gas chamber) "I can't take the chance." A spineless person would not need to deliberate....

One might argue that not all "hard choices" are examples of duress. A doctor could offer to sell vital treatment for a large sum. After the patient has been cured it will hardly do for her to claim that she has been the victim of duress. The doctor may have forced the patient to choose between a certain financial loss and the risk of death. But surely doctors are not like gunmen.

Two important points need to be made in response to this objection. First, the doctor is not, one assumes, responsible for the diseased condition of the patient. The patient would be facing death even if she had never met the doctor. But this is not true in the case of the gunman, where both impositions are his work. And in this respect the prosecutor offering a plea bargain in a criminal case is like the gunman rather than like the doctor. For the state forces a choice between adverse consequences that it imposes. And, of course, one cannot say that in the defendant's wrongdoing he has brought his dreadful dilemma upon himself. To do so would be to ignore the good reasons there are for the presumption of innocence in dispositive criminal proceedings.

Second, our laws do not prohibit doctors from applying their healing skills to maximize their own wealth. They are free to contract to perform services in return for a fee. But our laws do severely restrict the state in its prosecution of criminal defendants. Those who framed our constitution were well aware of the great potential for abuse that the criminal law affords. Much of the Constitution (especially the Bill of Rights) checks the activity of the state in this area. In particular, the Fifth Amendment provides that no person "shall be compelled in any criminal case to be a witness against himself." If I am right in judging that defendants like Alford and Davis's client do not act freely in pleading guilty to the facts of their cases, that the forced choice of the prosecutor may be as coercive as the forced choice of the gunman, that a defendant may be compelled to speak against himself (or herself) by a prosecutor's discretion inducing him to plead guilty, then given the apparent constitutional prohibition of such compulsion, the prosecutor acts wrongfully in compelling such pleas. And in this manner it may be that the last element of duress, wrongfulness, can be established. But it is not my purpose here to establish the unconstitutionality of plea bargaining, for it is not necessary to reach to unconstitutionality to grasp the wrongfulness of that institution. One need only reflect upon what justice amounts to in our system of criminal law. This is the task I will take up in the final section of this paper.

2

Not too long ago plea bargaining was an officially prohibited practice. Court procedures were followed to ensure that no concessions had been given to defendants in exchange for guilty pleas. But gradually it became widely known that these procedures had become charades of perjury, shysterism, and bad faith involving judges, prosecutors, defense attorneys and defendants. This was scandalous. But rather than cleaning up the practice in order to square it with the rules, the rules were changed in order to bring them in line with the practice....

Without going deeply into detail, I believe that it can be asserted without controversy that the

liberal-democratic approach to criminal justice—and in particular the American criminal justice system—is an institutionalization of two principles. The first principle refers to the intrinsic point of systems of criminal justice.

A. ***Those (and only those) individuals who are clearly guilty of certain serious specified wrongdoings deserve an officially administered punishment which is proportional to their wrongdoing.*** In the United States it is possible to see this principle underlying the activities of legislators specifying and grading wrongdoings which are serious enough to warrant criminalization and, further, determining the range of punishment appropriate to each offense; the activities of policemen and prosecutors bringing to trial those who are suspected of having committed such wrongdoings; the activities of jurors determining if defendants are guilty beyond a reasonable doubt; the activities of defense attorneys insuring that relevant facts in defendants' favor are brought out at trial; the activities of judges seeing to it that proceedings are fair and that those who are convicted receive the punishment they deserve; and the activities of probation officers, parole officers, and prison personnel executing the sentences of the courts. All of these people play a part in bringing the guilty to justice.

But in liberal-democratic societies not everything is done to accomplish this end. A second principle makes reference to the limits placed upon the power of the state to identify and punish the guilty.

B. ***Certain basic liberties shall not be violated in bringing the guilty to justice.*** This second principle can be seen to underlie the constellation of constitutional checks on the activities of virtually every person playing a role in the administration of the criminal justice system.

Each of these principles is related to a distinctive type of injustice that can occur in the context of criminal law. An injustice can occur in the outcome of the criminal justice procedure. That is, an innocent defendant may be convicted and punished, or a guilty defendant may be acquitted or, if convicted, he or she may receive more or less punishment than is deserved. Because these injustices occur in the meting out of punishment to defendants who are being processed by the system, we can refer to them as internal injustices. They are violations of the first principle. On the other hand, there is a type of injustice which occurs when basic liberties are violated in the operation of the criminal justice system. It may be true that Star Chamber proceedings, torture, hostages, bills of attainder, dragnet arrests, unchecked searches, *ex post facto* laws, unlimited invasions of privacy, and an arsenal of other measures could be employed to bring more of the guilty to justice. But these steps lead to a dystopia where our most terrifying nightmares can come true. However we limit the activity of the criminal justice system in the interest of basic liberty, that limit can be overstepped. We can call such infringements upon basic liberties external injustices. They are violations of the second principle. If, for example, what I have suggested in the previous section is correct, then plea bargaining can bring about an external injustice with respect to a basic liberty secured by the Fifth Amendment. The remainder of this section will be concerned with internal injustice or violations of the first principle.

It is necessary to draw a further distinction between aberrational and systemic injustice. It may very well be that in the best criminal justice system that we are capable of devising human limitations will result in some aberrational injustice. Judges, jurors, lawyers, and legislators with the best of intentions may make errors in judgment that result in mistakes in the administration of punishment. But despite the knowledge that an unknown percentage of all dispositions of criminal cases are, to some extent, miscarriages of justice, it may still be reasonable to believe that a certain system of criminal justice is well calculated to avoid such results within the limits referred to by the second principle. We can refer to these incorrect outcomes of a sound system of criminal justice as instances of aberrational injustice. In contrast, instances of systemic injustice are those that result from structural flaws in the criminal justice system itself. Here incorrect outcomes in the operations of the system are not the result of human error. Rather, the system itself is not well calculated to avoid injustice. What would be instances of aberrational injustice in a sound system are not aberrations in an unsound system: they are a standard result.

. . . [S]ystemic injustice in the context of criminal law is a much more serious matter than aberrational injustice. It should not be forgotten that the criminal sanction is the most severe imposition that the state can visit upon one of its citizens. While it is possible to tolerate occasional error in a sound system, systematic carelessness in the administration of punishment is negligence of the highest order.

With this framework in mind, let us look at a particular instance of plea bargaining recently described by a legal aid defense attorney. Ted Alston has been charged with armed robbery. Let us assume that persons who have committed armed robbery (in the way Alston is accused of having committed it) deserve five to seven years of prison. Alston's attorney sets out the options for him: "I told Alston it was possible, perhaps even probable, that if he went to trial he would be convicted and get a prison term of perhaps five to seven years. On the other hand, if he agreed to plead guilty to a low-grade felony, he would get a probationary sentence and not go to prison. The choice was his." Let us assume that Alston accepts the terms of the bargain and pleads guilty to a lesser offense. If Alston did commit the armed robbery, there is a violation of the first principle in that he receives far less punishment than he deserves. On the other hand, if Alston did not commit the armed robbery, there is still a violation of the first principle in that he is both convicted of and punished for a crime that he did not commit, a crime that no one seriously believes to be his distinctive wrongdoing. It is of course possible that while Alston did not commit the armed robbery, he did commit the lesser offense. But though justice would be done here, it would be an accident. Such a serendipitous result is a certain sign that what we have here is systemic injustice.

If we assume that legislatures approximate the correct range of punishment for each offense, that judges fairly sentence those who are convicted by juries, and that prosecutors reasonably charge defendants, then, barring accidents, justice will *never* be the outcome of the plea-bargaining procedure: the defendant who "cops a plea" will never receive the punishment which is deserved. Of course legislatures can set punishments too high, judges can oversentence those who are convicted by juries, and prosecutors can overcharge defendants. In these cases the guilty can receive the punishment they deserve through plea bargaining. But in these cases we compensate for one injustice by introducing others that unfairly jeopardize the innocent and those that demand trials.

In contrast to plea bargaining, the disposition of criminal cases by jury trial seems well calculated to avoid internal injustices even if these may sometimes occur. Where participants take their responsibilities seriously we have good reason to believe that the outcomes is just, even when this may not be so. In contrast, with plea bargaining we have no reason to believe that the outcome is just even when it is.

I think that the appeal that plea bargaining has is rooted in our attitude toward bargains in general. Where both parties are satisfied with the terms of an agreement, it is improper to interfere. Generally speaking, prosecutors and defendants are pleased with the advantages they gain by negotiating a plea. And courts, which gain as well, are reluctant to vacate negotiated pleas where only "proper" inducements have been applied and where promises have been understood and kept. Such judicial neutrality may be commendable where entitlements are being exchanged. But the criminal justice system is not such a context. Rather it is one in which persons are justly given, not what they have bargained for, but what they deserve, irrespective of their bargaining position.

To appreciate this, let us consider another context in which desert plays a familiar role; the assignment of grades in an academic setting. Imagine a "grade bargain" negotiated between a grade-conscious student and a harried instructor. A term paper has been submitted and, after glancing at the first page, the instructor says that if he were to read the paper carefully, applying his usually rigid standards, he would probably decide to give the paper a grade of D. But if the student were to waive his right to a careful reading and conscientious critique, the instructor would agree to a grade of B. The grade-point average being more important to him than either education or justice in grading, the student happily accepts the B, and the instructor enjoys a reduced workload.

One strains to imagine legislators and administrators commending the practice of grade bargaining because it permits more students to be processed by fewer instructors. Teachers can be freed from the

burden of having to read and to criticize every paper. One struggles to envision academicians arguing for grade bargaining, suggesting that a quick assignment of a grade is a more effective influence on the behavior of students, urging that grade bargaining is necessary to the efficient functioning of the schools. There can be no doubt that students who have negotiated a grade are more likely to accept and to understand the verdict of the instructor. Moreover, in recognition of a student's help to the school (by waiving both the reading and the critique), it is proper for the instructor to be lenient. Finally, a quickly assigned grade enables the guidance personnel and the registrar to respond rapidly and appropriately to the student's situation.

What makes all of this laughable is what makes plea bargaining outrageous. For grades, like punishments, should be deserved. Justice in retribution, like justice in grading, does not require that the end result be acceptable to the parties. To reason that because the parties are satisfied the bargain should stand is to be seriously confused. For bargains are out of place in contexts where persons are to receive what they deserve. And the American courtroom, like the American classroom, should be such a context.

In this section, until now I have been attempting to show that plea bargaining is not well calculated to insure that those guilty of wrongdoing will receive the punishment they deserve. But a further point needs to be made. While the conviction of the innocent would be a problem in any system we might devise, it appears to be a greater problem under plea bargaining. With the jury system the guilt of the defendant must be established in an adversary proceeding and it must be established beyond a reasonable doubt to each of twelve jurors. This is very staunch protection against an aberrational conviction. But under plea bargaining the foundation for conviction need only include a factual basis for the plea (in the opinion of the judge) and the guilty plea itself. Considering the coercive nature of the circumstances surrounding the plea, it would be a mistake to attach much reliability to it. Indeed, as we have seen in *Alford,* guilty pleas are acceptable even when accompanied by a denial of guilt. . . . Now it is one thing to show to a judge that there are facts which support a plea of guilty and quite another to prove to twelve jurors in an adversary proceeding guilt beyond a reasonable doubt. Plea bargaining substantially erodes the standards for guilt and it is reasonable to assume that the sloppier we are in establishing guilt, the more likely it is that innocent persons will be convicted. So apart from having no reason whatever to believe that the guilty are receiving the punishment they deserve, we have far less reason to believe that the convicted are guilty in the first place than we would after a trial.

In its coercion of criminal defendants, in its abandonment of desert as the measure of punishment, and in its relaxation of the standards for conviction, plea bargaining falls short of the justice we expect of our legal system. I have no doubt that substantial changes will have to be made if the institution of plea bargaining is to be obliterated or even removed from its central position in the criminal justice system. No doubt we need more courts and more prosecutors. Perhaps ways can be found to streamline the jury trial procedure without sacrificing its virtues. Certainly it would help to decriminalize the host of victimless crimes—drunkenness and other drug offenses, illicit sex, gambling, and so on—in order to free resources for dealing with more serious wrongdoings. And perhaps crime itself can be reduced if we begin to attack seriously those social and economic injustices that have for too long sent their victims to our prisons in disproportionate numbers. In any case, if we are to expect our citizenry to respect the law, we must take care to insure that our legal institutions are worthy of that respect. I have tried to show that plea bargaining is not worthy, that we must seek a better way. Bargain justice does not become us.

REVIEW AND DISCUSSION QUESTIONS

1. Is it fair that defendants who plead guilty are given more lenient sentences than those who are found guilty by trial?
2. Do prosecutor and defendant have equal bargaining weight in plea negotiations? Does plea bargaining involve duress or just a "hard choice"? If duress, is there something unfair about this duress? Is Kipnis's

analogy between plea bargaining and a gunman demanding one's wallet an accurate comparison? Or is plea bargaining more like a doctor offering to sell a vital treatment or an auto mechanic offering to assist a stranded motorist for a large sum?

3. What two principles does our criminal justice system institutionalize, and how is each related to a distinctive type of injustice? What is the difference between aberrational and systemic injustice? On what grounds does Kipnis argue that plea bargaining involves systemic injustice? Do you agree?
4. Is Kipnis right when he claims that prosecutors act unjustly by not assuming that the person is innocent? Explain.

Convicting the Innocent

James McCloskey

Most people assume that it is very rare for an innocent person to be convicted of a crime and that such a miscarriage of justice would be an isolated aberration in an otherwise sound system. James McCloskey of Centurion Ministries in Princeton, New Jersey, argues to the contrary. He estimates that at least 10 percent of those convicted of serious and violent crimes are completely innocent, and in this essay, he examines seven major causes of wrongful convictions.

On most occasions when it has been discovered that the wrong person was convicted for another's crime, the local law enforcement community, if it has commented at all, has assured the public that such instances are indeed rare and isolated aberrations of a criminal justice system that bats nearly 1,000 percent in convicting the guilty and acquitting the innocent. And this view is shared, I think, not only by the vast majority of the public but also by almost all of the professionals (lawyers and judges) whose work comes together to produce the results.

I realize that I am a voice crying in the wilderness, but I believe that the innocent are convicted far more frequently than the public cares to believe, and far more frequently than those who operate the system dare to believe. An innocent person in prison, in my view, is about as rare as a pigeon in the park. The primary purpose of this article is to delineate why and how I have come to believe that this phenomenon of the "convicted innocent" is so alarmingly widespread in the United States. Although no one has any real idea of what proportion it has reached, it is my perception that at least 10 percent of those convicted of serious and violent crimes are completely innocent. Those whose business it is to convict or to defend would more than likely concede to such mistakes occurring in only 1 percent of cases, if that. Regardless of where the reader places his estimate, these percentages, when converted into absolute numbers, tell us that thousands and even tens of thousands of innocent people languish in prisons across the nation.

Allow me to outline briefly the ground of experience on which I stand and speak. For the past eight years I have been working full time on behalf of the innocent in prison. To date, the nonprofit organization I founded to do this work has freed and vindicated three innocent lifers in New Jersey. Another, on Texas's death row, has been declared "innocent" by a specially appointed evidentiary hearing judge, who has recommended a new trial to Texas's highest court. Currently we are working on ten cases across the country (New Jersey, Pennsylvania, Virginia, Louisiana, Texas, and California). We have received well over 1,000 requests for assistance and have developed extensive files on more than 500 of these requests, which come to us daily from every state of the nation from those who have been convicted, or from their advocates, proclaiming their innocence. We serve as active advisors on many of those cases.

Besides being innocent and serving life or death sentences, our beneficiaries have lost their legal appeals. Their freedom can be secured only by developing new evidence sufficient to earn a retrial. This new evidence must materially demonstrate either that the person is not guilty or that the key state witnesses lied in critical areas of their testimony. We are not lawyers. We are concerned only with whether the person is in fact completely not guilty in that he or she had nothing whatsoever to do with the crime. When we enter the case it is usually five to fifteen years after the conviction. Our sole focus is to reexamine the factual foundation of the conviction—to conduct an exhaustive investigation of the cast of characters and the circumstances in the case, however long that might take. . . .

APPELLATE RELIEF FOR THE CONVICTED INNOCENT

As all lawyers and jurists know, but most lay people do not, innocence or guilt is irrelevant when seeking redress in the appellate courts. As the noted attorney F. Lee Bailey observed, "Appellate courts have only one function, and that is to correct legal mistakes of a serious nature made by a judge at a lower level. Should a jury have erred by believing a lying witness, or by drawing an attractive but misleading inference, there is nothing to appeal." So, if the imprisoned innocent person is unable to persuade the appellate judges of any legal errors at trial, and generally he cannot, even though he suffered the ultimate trial error, he has no recourse. Nothing can be done legally to free him unless new evidence somehow surfaces that impeaches the validity of the conviction. Commonly, the incarcerated innocent are rubber-stamped into oblivion throughout the appeals process, both at the state and at the federal level. . . .

Once he is convicted, no one in whose hands his life is placed (his lawyer and the appellate judges) either believes him or is concerned about his innocence or guilt. It is no longer an issue of relevance. The only question remaining that is important or material is whether he "legally" received a fair trial, not whether the trial yielded a result that was factually accurate. Appellate attorneys are not expected to, nor do they have the time, inclination, and resources to, initiate an investigation designed to unearth new evidence that goes to the question of a false conviction. Such an effort is simply beyond the scope of their thinking and beyond the realm of their professional responsibility. It is a rare attorney indeed who would dare go before any American appellate court and attempt to win a retrial for his client based on his innocence. That's like asking an actor in a Shakespearian tragedy to go on stage and pretend it's a comedy. It is simply not done.

CAUSES OF WRONGFUL CONVICTION

But enough of this post-conviction appellate talk. That's putting the cart before the horse. Let's return to the trial and discuss those elements that commonly combine to convict the innocent. Let me state at the outset that each of these ingredients is systemic and not peculiar to one part of the country or one type of case. We see these elements as constant themes or patterns informing the cases that cross our desks. They are the seeds that sow wrongful convictions. After one has reflected on them individually and as a whole, it becomes readily apparent, I think, how easy it is and how real the potential is in every courthouse in America for wrongful convictions to take place.

(A) Presumption of Guilt

The first factor I would like to consider is the "presumption-of-innocence" principle. Although we would all like to believe that a defendant is truly considered innocent by those who represent and judge him, this is just not so. Once accusations have matured through the system to the point at which the accused is actually brought to trial, is it not the tendency of human nature to suspect deep down or even believe that the defendant probably did it? Most people are inclined to believe that where there is smoke, there is fire. This applies to professional and lay people alike albeit for different reasons perhaps.

The innate inclinations of the average American law-abiding citizen whose jury experience is that

person's first exposure to the criminal justice system is to think that law enforcement people have earnestly investigated the case and surely would not bring someone to trial unless they had bona fide evidence against the person. That is a strong barrier and a heavy burden for the defense to overcome. And how about judges and defense lawyers? These professionals, like members of any profession, have a natural tendency to become somewhat cynical and callous with time. After all, isn't it true that the great majority of the defendants who have paraded before them in the past have been guilty? Why should this case be any different? As far as defense attorneys are concerned, if they really believe in their clients' innocence, why is that in so many instances they are quick to urge them to take a plea for a lesser sentence than they would get with a trial conviction? So, by the time a person is in the trial docket, the system (including the media) has already tarnished him with its multitude of prejudices, which of course, would all be denied by those who entertain such prejudices.

(B) Perjury by Police

Another reason for widespread perversions of justice is the pervasiveness of perjury. The recent District Attorney of Philadelphia once said, "In almost any factual hearing or trial, someone is committing perjury; and if we investigate all of those things, literally we would be doing nothing but prosecuting perjury cases." If he is guilty, the defendant and his supporters would lie to save his skin and keep him from going to prison. That is assumed and even expected by the jury and the judge. But what would surprise and even shock most jury members is the extent to which police officers lie on the stand to reinforce the prosecution and not jeopardize their own standing within their own particular law enforcement community. The words of one twenty-five-year veteran senior officer of a northern New Jersey police force still ring in my ears: "They [the defense] lie, so we [police] lie. I don't know one of my fellow officers who hasn't lied under oath." Not too long ago a prominent New York judge, when asked if perjury by police was a problem, responded, "Oh, sure, cops often lie on the stand."

(C) False Witnesses for the Prosecution

What is more, not only do law officers frequently lie, but the primary witnesses for the prosecution often commit perjury for the state, and do so under the subtle guidance of the prosecutor. Inveterately, common criminals who are in deep trouble themselves with the same prosecutor's office or local police authority are employed as star state witnesses. In exchange for their false testimony, their own charges are dismissed, or they are given non-custodial or greatly reduced prison sentences. In other words a secret deal is struck whereby the witness is paid for his fabricated testimony with that most precious of all commodities—freedom!

Such witnesses are usually brought forward by the state to say either that the defendant confessed the crime to them or that they saw the defendant near the crime scene shortly before it happened, or they saw him flee the scene of the crime as it was occurring. If I have seen one, I have seen a hundred "jailhouse confessions" spring open the prison doors for the witness who will tell a jury on behalf of the state that the defendant confessed the crime to him while they shared the same cell or tier. When the state needs important help, it goes to its bullpen, the local county jail, and brings in one of the many ace relievers housed there to put out the fire. As several of these "jailhouse priests" have told me, "It's a matter of survival: either I go away or he [the defendant] goes away, and I'm not goin'." Jailhouse confessions are a total perversion of the truth-seeking process. Amazingly enough, they are a highly effective prosecutorial means to a conviction. Part and parcel of a jailhouse confession is the witness lying to the jury when he assures them that he expects nothing in return for his testimony, that he is willing to swallow whatever pill he must for his own crimes.

(D) Prosecutorial Misconduct

The right decision by a jury depends largely on prosecutorial integrity and proper use of prosecutorial power. If law enforcement officers, in their zeal to win and convict, manipulate or intimidate witnesses into false testimony, or suppress evidence that impeaches the prosecution's own witnesses or even goes to the defendant's innocence, then the

chances of an accurate jury verdict are greatly diminished. Sadly, we see this far too often. It is frightening how easily people respond to pressure or threats of trouble by the authorities of the law. Our insecurities and fears as well as our desires to please those who can punish us allow all of us to be far more malleable than we like to think.

Few of us have the inner strength we think we have to resist such overreaching by the law. This applies to mainline citizenry as well as to those living on the margins. However, the underclasses are particularly vulnerable and susceptible to police pressure because they are powerless; and both they and the police know it. A few examples will illustrate.

In 1981 three white high school janitors were threatened by the Texas Rangers into testifying that they had seen Clarence Brandley, their black custodial supervisor, walking into the restroom area of the high school where the victim had entered only minutes before she had disappeared. Brandley was convicted and sentenced to death based on the inferential testimony that since he was the last person seen near her, then he must have killed her. Eight years later Brandley was exonerated by the judge who conducted his evidentiary hearing when one of these janitors came forward and told how they had lied in implicating Brandley because of coercion by the investigating law officer.

On the eve of the Rene Santana trial in Newark, New Jersey, which was a year and a half after the crime, the prosecutors produced a surprise "eyewitness" who said he saw Mr. Santana flee the scene of the crime. A decade later that same witness visited Mr. Santana at New Jersey's Rahway State Prison and asked for his forgiveness after admitting to him that he had concocted the "eyewitness" testimony in response to intense pressure from the prosecutor's investigator. Since this "eyewitness" was from Trujillo's Dominican Republic police state, his innate fear of the police made him vulnerable to such police coercion.

Or how about the Wingo case in white, rural northwestern Louisiana? Wingo's common-law wife came forward on the eve of his execution and admitted that she had lied at his trial five years earlier because the deputy sheriff had threatened to put her in jail and forever separate her from her children unless she regurgitated at trial what he wanted her to say.

And in the Terry McCracken case in the suburbs of Philadelphia, a fellow high school student of the caucasian McCracken testified that he saw McCracken flee the convenience store moments after a customer was shot to death during the course of a robbery. The teenager was induced to manufacture this false eyewitness account after three visits to the police station. Among the evidence that vindicates McCracken are the confessions by the real robber/killers. So, you see, it not only can happen anywhere, it does happen everywhere; and it does happen to all different people, regardless of race and background.

Another common trait of wrongful convictions is the prosecutor's habit of suppressing or withholding evidence which he is obliged to provide to the defendant in the interests of justice and fairness. Clarence Darrow was right when he said, "A courtroom is not a place where truth and innocence inevitably triumph; it is only an arena where contending lawyers fight not for justice but to win." And so many times this hidden information is not only "favorable" to the defendant but it clears him. In Philadelphia's Miguel Rivera case the district attorney withheld the fact that two shopkeepers had seen the defendant outside their shop when the art museum murder was actually in progress. And in the Gordon Marsh case near Baltimore, Maryland, the state failed to tell the defendant that its main witness against him was in jail when she said she saw him running from the murder scene. One has to wonder what the primary objective of prosecutors is. Is it to convict, regardless of the factual truth, or is it to pursue justice?

The prosecution is the "house" in the criminal justice system's game of poker. The cards are his, and he deals them. He decides whom and what to charge for crimes, and if there will be a trial or whether a plea is acceptable. He dominates. Unfortunately, his power is virtually unchecked because he is practically immune from punishment for offenses, no matter how flagrant or miscreant. According to many state and federal courts, prosecutorial misbehavior occurs with "disturbing frequency." When the "house" cheats, the innocent lose. . . .

It is human nature to resist any information that indicates that we have made a grievous mistake. This

is particularly true of prosecutors when presented with new evidence that impeaches a conviction and goes to the innocence of a person convicted by their office at a prior time, whether it occurred four months or forty years before. Not only are they coldly unresponsive to such indications but they quickly act to suppress or stamp them out. New evidence usually comes in the form of a state witness who, plagued with a guilty conscience, admits that he lied at the trial; or from a person completely new to the case who comes forward with his exculpatory knowledge. Without exception, in my experience, the prosecutor's office will treat that person with total contempt in its usually successful attempt to force the person to retreat into silence. If that doesn't work, it will dismiss such testimony as somehow undeserving of any credibility and blithely ignore it. This prosecutorial impishness reminds me of a little boy holding his hands to his ears on hearing an unpleasant sound.

The Joyce Ann Brown case is a poignant illustration of this kind of prosecutorial posturing. One year after Joyce's 1980 conviction for being one of two black women who had robbed a Dallas, Texas furrier and killed one of the proprietors, the admitted shooter was captured and pleaded guilty while accepting a life sentence. She also told her attorney that the district attorney had convicted the wrong woman (Joyce Brown) as her partner in the crime. She had never known or even heard of that Joyce Brown. With the district attorney fighting her with all of his might, Joyce sits in prison to this day trying to win a retrial as we try to develop new evidence on her behalf.

(E) Shoddy Police Work

The police work of investigating crimes, when done correctly and thoroughly, is indeed a noble profession. Law and order are essential to a cohesive and just society. Because police work is fraught with so many different kinds of pressures, it is rather easy for an investigation to go awry. The high volume of violent crime plagues every urban police department. Skilled detectives are few, and their caseloads are overwhelming. The "burnout" syndrome is a well-documented reality within police ranks. Interdepartmental politics and the bureaucracy stifle initiative and energy. The pressure to "solve" a case is intensely felt by the line detective and comes both from his superiors and the community and from his own ambitious need for recognition and advancement. If today's climate of "burn or bury" them puts more pressure on the detective to resolve, it also gives him more license to do so by whatever means.

Too often, as a result of the above factors, police officers take the easy way out. Once they come to suspect someone as the culprit, and this often occurs early within the investigation and is based on rather flimsy circumstantial information, then the investigation blindly focuses in on that adopted "target." Crucial pieces of evidence are overlooked and disregarded. Some witnesses are not interviewed who should be, while others are seduced or coerced into telling the police what they want to hear. Evidence or information that does not fit the suspect or the prevailing theory of the crime is dismissed as not material or is changed to implicate the suspect. Good old-fashioned legwork is replaced by expediency and shortcuts. Coercive confessions are extracted and solid leads are ignored.

Before too long, momentum has gathered, and the "project" now is to put it on the suspect. Any information that points to the suspect, no matter how spuriously secured, is somehow obtained; and anything that points away from him is ridiculed and twisted into nothingness. The task is made much easier if the suspect has a police record because he should be "taken off the streets" anyhow. That kind of person is not only a prime suspect but also a prime scapegoat. An example of this is Clarence Brandley, who was mentioned earlier. He was arrested in late August four days after the crime and on the weekend before school was to begin. The high school where the rape and murder took place was flooded with telephone calls by scared parents who refused to send their children to school until the murderer was caught. The arrest of Brandley calmed the community, and school started as scheduled. It was after Brandley's arrest that the investigation then spent five hundred hours building the case against him.

(F) Incompetent Defense Counsel

The wrongly convicted invariably find themselves between the rock of police/prosecutorial misconduct and the hard place of an incompetent and irresponsible defense attorney. While the

correct decision by a jury hinges on a fair prosecution, it also depends on dedicated and skilled defendant lawyering. And there is such a paucity of the latter. Not only are there very few highly competent defense lawyers, but there are very few criminal defense lawyers, period. They are rapidly becoming an extinct species.

The current Attorney General of New Jersey not too long ago told the New Jersey State Bar Association that finding quality private defense attorneys "may be the most crying need that we have." He also told this same assemblage that unless there is an adequate number of well-trained private defense lawyers, there will be little hope for justice. Of the 30,000 lawyers in New Jersey, the number of those doing primarily criminal defense work is only in the hundreds. At this same conference the First Assistant Attorney General pointed out that 85 percent of New Jersey's criminal cases are handled by the public defender system; and he wondered if there would be a private defense bar [in the future].

This means, of course, that 85 percent of those charged with a crime cannot afford an attorney, so they are forced to use the public defender system. As competent as New Jersey's full-time salaried public defenders generally are, their resources (budget and people) are vastly inadequate and are dwarfed by those of their adversaries (the local prosecutor's office). Moreover, they are so overwhelmed by the sheer volume of caseload that no defender can give quality attention to any one of his cases, let alone all of them. So, in response to this shortage, public defender cases are farmed out to "pooled" attorneys, who are paid a pittance relative to what they earn from other clients who retain them privately.

The experience of these pooled attorneys in criminal matters is often limited and scanty. In addition, they do not bring to their new-found indigent client the desired level of heart and enthusiasm for their cases. All of these conditions leave the defendant with an attorney somewhat lacking in will, effort, resources, and experience. Thus, the defendant goes to trial with two strikes against him.

What we have discovered as a common theme among those whose cases we have studied from all over the country is that their trial attorney, whether from the public domain or privately retained, undertakes his work with an appalling lack of assiduity. Communication with the defendant is almost nonexistent. When it does take place, it is carried on in a hurried, callous, and dismissive manner. Attempts at discovery are made perfunctorily. Prosecutors are not pressed for this material. Investigation is shallow and narrow, if conducted at all. Preparation meets minimal standards. And advocacy at trial is weak. Cross-examination is superficial and tentative.

Physical evidence is left untested, and forensic experts are not called to rebut whatever scientific evidence the state introduces through its criminalists. I cannot help thinking of the Nate Walker case, where, at Nate's 1976 trial for rape and kidnapping, the doctor who examined the victim the night of her ordeal testified that he found semen in her vaginal cavity. Walker's privately retained attorney had no questions for the doctor when it came time for cross-examination, nor did he even ask anyone to test the vaginal semen for blood type. Twelve years later, that test was performed at our request, and Walker was exonerated and immediately freed. . . .

(G) Nature of Convicting Evidence

The unschooled public largely and erroneously believes that convictions are mostly obtained through the use of one form of tangible evidence or another. This naive impression is shaped by watching too many TV shows like Perry Mason or Matlock. The reality is that in most criminal trials the verdict more often than not hinges on whose witnesses—the state's or defendant's—the jury chooses to believe. It boils down to a matter of credibility. There is no "smoking gun" scientific evidence that clearly points to the defendant. This puts an extremely heavy burden on the jury. It must somehow ferret out and piece together the truth from substantially inconsistent and contradictory testimony between and within each side. The jury is forced to make one subjective call after another in deciding whom to believe and what inferences to draw from conflicting statements.

For example, how can a jury accept a victim's positive identification at trial of the defendant as her assailant when she had previously described her attacker in physical terms that were very different from the actual physical characteristics of the

defendant, or when the defense has presented documented information that precludes the defendant from being the assaulter? Several cases come to mind. Boy was convicted of robbing a convenience store in Georgia. The clerk initially told the police that since she was 5 feet 3 inches, was standing on a 3-inch platform, and had direct eye contact with the robber, he must have been about 5 feet 6 inches tall. Boy is 6 feet 5 inches tall. Four teenage girls identified Russell Burton as their rapist on a particular day in Arkansas. Burton introduced evidence that on that day his penis was badly blistered from an operation two days before for removal of a wart. And a Virginia woman was certain that Edward Honaker was her rapist even though her rapist had left semen within her, and Honaker had had a vasectomy well in advance of the assault.

Criminal prosecutions that primarily or exclusively depend on the victim's identification of the defendant as the perpetrator must be viewed with some skepticism unless solid corroborating evidence is also introduced. Traumatized by a crime as it occurs, the victim frequently is looking but not seeing. Victims are extremely vulnerable and can easily be led by the police, through unduly suggestive techniques, into identifying a particular person. The victim in Nate Walker's case, for example, was with her abductor/rapist for two and a half hours with ample opportunity to clearly view him. She told the jury without hesitation eighteen months later that "he's the man." Nate had an ironclad alibi. The jury struggled for several days but in the end came in with a guilty verdict. An mentioned earlier, he was scientifically vindicated twelve years later.

When juries are confronted with a choice between a victim's ringing declaration that "that's the man" and solid evidence that "it couldn't be him," they usually cast their lot with the victim. I suggest that this can be a very dangerous tendency and practice. And this is particularly so when identification crosses racial lines, that is, when a white victim says it was that black person. Future jurors should be aware that identifications can be very unreliable forms of evidence.

Another type of evidence that can be misleading and even confusing to jurors is that offered by laboratory scientists. Results of laboratory tests that are presented by the forensic scientists are not always what they appear to be, although they strongly influence jury decisions. A recent New York Times article pointed out that there is a "growing concern about the professionalism and impartiality of the laboratory scientists whose testimony in court can often mean conviction or acquittal." This article went on to say that the work of forensic technicians in police crime laboratories is plagued by uneven training and questionable objectivity.

We share this mounting concern because we see instance after instance where the prosecutor's crime laboratory experts cross the line from science to advocacy. They exaggerate the results of their analysis of hairs, fibers, blood, or semen in such a manner that it is absolutely devastating to the defendant. To put the defendants at a further disadvantage, the defense attorneys do not educate themselves in the forensic science in question, and therefore conduct a weak cross-examination. Also, in many cases, the defense does not call in its own forensic experts, whose testimony in numerous instances could severely damage the state's scientific analysis.

One case profoundly reflects this common cause of numerous unjust convictions. Roger Coleman sits on Virginia's death row today primarily because the Commonwealth's Bureau of Forensic Science expert testified that the two foreign pubic hairs found on the murdered victim were "consistent" with Mr. Coleman's, and that it was "unlikely" that these hairs came from someone other than Mr. Coleman. The defense offered nothing in rebuttal, so this testimony stood unchallenged. In a post-conviction hearing Mr. Coleman's new lawyer introduced the testimony of a forensic hair specialist who had twenty-five years of experience with the F.B.I. He testified that "it is improper to conclude that it is likely that hairs came from a particular person simply because they are consistent with that person's hair because hairs belonging to different people are often consistent with each other, especially pubic hairs."

Another problem that we continually observe within the realm of forensic evidence is the phenomenon of lost and untested physical evidence. Often, especially in cases up to the early 1980s, the specimens that have the potential to exclude the defendant have not been tested and eventually get misplaced. At best this is gross negligence on the part of both the police technician and the defense attorney in not ensuring that the tests be done.

CONCLUSION

. . . My contention is that at least 10 percent of those convicted for serious, violent crimes are incorrectly convicted because some combination of the trial infirmities described in this article results in mistaken jury determinations.

Everyone will agree that the system is not perfect, but the real question is this: To what extent do its imperfections prevail? I contend that for all the reasons detailed above the system is a far leakier cistern than any among us has ever imagined. Untold numbers of innocents have tumbled into the dark pit of prison. Some of them have eventually gained their freedom, but a majority remain buried in prison, completely forsaken and forgotten by the outside world.

REVIEW AND DISCUSSION QUESTIONS

1. What leads McCloskey to say that "innocence or guilt is irrelevant when seeking redress in the appellate courts"?
2. McCloskey identifies the following as causes of wrongful convictions: presumption of guilt, perjury by police, false witnesses for the prosecution, prosecutorial misconduct, shoddy police work, incompetent defense counsel, and the nature of conflicting evidence. Assess what he has to say about each of these factors. Which do you think are the most important?
3. Are you persuaded by McCloskey that many innocent people are wrongfully convicted? If he is right, why is the contrary so widely believed?
4. What can be done to reduce or eliminate wrongful convictions? Which of the causes identified by McCloskey could be most easily corrected? Which would be the most difficult to remedy?
5. Assume that McCloskey is correct about the number of wrongful convictions. What implications does this have for our adversary system? Is that system one of the causes of wrongful convictions, or is it part of the solution to that problem?

The Dirty Harry Problem

Carl B. Klockars

Police officers sometimes face morally difficult situations. In this essay, Carl Klockars uses the Clint Eastwood film *Dirty Harry* to illustrate an important and troubling dilemma of police work: whether a morally good end, such as rescuing a kidnapping victim, can justify the use of morally "dirty" means, such as using force to extort a confession from a suspect. Klockars explores the factors that make the Dirty Harry problem a genuine moral dilemma and critically examines three possible responses to it, none of which succeeds in fully eliminating the dilemma.

When and to what extent does the morally good end warrant or justify an ethically, politically, or legally dangerous means for its achievement? This is a very old question for philosophers. . . . I know of not a single contribution to the criminological or sociological literature on policing which raises it explicitly and examines its implications. This is the case in spite of the fact that there is considerable evidence to suggest that it is not only an ineluctable part of police work, but a moral problem with which police themselves are quite familiar. There are, I believe, a number of good reasons why social scientists have avoided or neglected what I like to call the Dirty Harry problem in policing, not the

least of which is that it is insoluble. However, a great deal can be learned about police work by examining some failed solutions, three of which I consider in the following pages. First, though, it is necessary to explain what a Dirty Harry problem is and what it is about it that makes it so problematic.

THE DIRTY HARRY PROBLEM

The Dirty Harry problem draws its name from the 1971 Warner Brothers film *Dirty Harry* and its chief protagonist, antihero Inspector Harry "Dirty Harry" Callahan. The film features a number of events which dramatize the Dirty Harry problem in different ways, but the one which does so most explicitly and most completely places Harry in the following situation. A 14-year-old girl has been kidnapped and is being held captive by a psychopathic killer. The killer, "Scorpio," who has already struck twice, demands $200,000 ransom to release the girl, who is buried with just enough oxygen to keep her alive for a few hours. Harry gets the job of delivering the ransom and, after enormous exertion, finally meets Scorpio. At their meeting Scorpio decides to renege on his bargain, let the girl die, and kill Harry. Harry manages to stab Scorpio in the leg before he does so, but not before Scorpio seriously wounds Harry's partner, an inexperienced, idealistic, slightly ethnic, former sociology major.

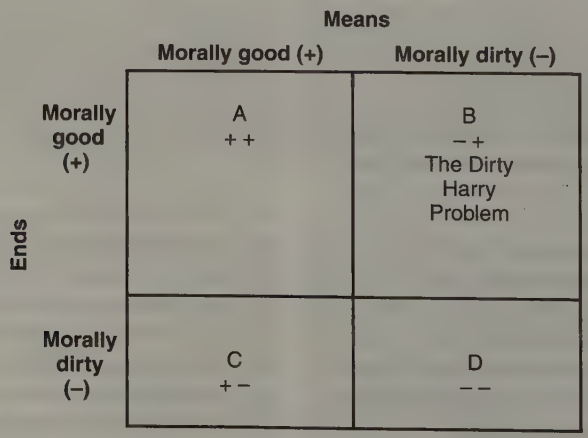
Scorpio escapes, but Harry manages to track him down through the clinic where he was treated for his wounded leg. After learning that Scorpio lives on the grounds of a nearby football stadium, Harry breaks into his apartment, finds guns and other evidence of his guilt, and finally confronts Scorpio on the 50-yard line, where Harry shoots him in the leg as he is trying to escape. Standing over Scorpio, Harry demands to know where the girl is buried. Scorpio refuses to disclose her location, demanding his rights to a lawyer. As the camera draws back from the scene Harry stands on Scorpio's bullet-mangled leg to torture a confession of the girl's location from him.

As it turns out, the girl is already dead and Scorpio must be set free. Neither the gun found in the illegal search, nor the confession Harry extorted, nor any of its fruits—including the girl's body—would be admissible in court.

The preceding scene, the heart of *Dirty Harry,* raises a number of issues of far-reaching significance for the sociology of the police, the first of which will now be discussed.

THE DIRTY HARRY PROBLEM I: THE END OF INNOCENCE

As we have phrased it previously, the Dirty Harry problem asks when and to what extent does the morally good end warrant or justify an ethically, politically, or legally dangerous means to its achievement? In itself, this question assumes the possibility of a genuine moral dilemma and posits its existence in a means-ends arrangement which may be expressed schematically as follows:



It is important to specify clearly the terms of the Dirty Harry problem not only to show that it must involve the juxtaposition of good ends and dirty means, but also to show what must be proven to demonstrate that a Dirty Harry problem exists. If one could show, for example, that box B is always empirically empty or that in any given case the terms of the situation are better read in some other means-ends arrangement, Dirty Harry problems vanish. At this first level, however, I suspect that no one could exclude the core scene of *Dirty Harry* from the class of Dirty Harry problems. There is no question that saving the life of an innocent victim of kidnapping is a "good" thing nor that grinding the bullet-mangled leg of Scorpio to extort a confession from him is "dirty."[1]

There is, in addition, a second level of criteria of an empirical and epistemological nature that must be met before a Dirty Harry problem actually comes into being. They involve the connection between the dirty act and the good end. Principally, what must be known and, importantly, known before the dirty act is committed, is that it will result in the achievement of the good end. In any absolute sense this is, of course, impossible to know, in that no acts are ever completely certain in their consequences. Thus the question is always a matter of probabilities. But it is helpful to break those probabilities into classes which attach to various subcategories of the overall question. In the given case, this level of problem would seem to require that three questions be satisfied, though not all with the same level of certainty.

In *Dirty Harry,* the first question is, Is Scorpio able to provide the information Dirty Harry seeks? It is an epistemological question about which, in *Dirty Harry,* we are absolutely certain. Harry met Scorpio at the time of the ransom exchange. Not only did he admit the kidnapping at that time, but when he made the ransom demand, Scorpio sent one of the girl's teeth and a description of her clothing and underwear to leave no doubt about the existence of his victim.

Second, we must know there are means, dirty means and nothing other than dirty means, which are likely to achieve the good end. One can, of course, never be sure that one is aware of or has considered all possible alternatives, but in *Dirty Harry* there would appear to be no reason for Scorpio in his rational self-interest to confess to the girl's location without being coerced to do so.

The third question which must be satisfied at this empirical and epistemological level concedes that dirty means are the only method which will be effective, but asks whether or not, in the end, they will be in vain. We know in *Dirty Harry* that they were, and Harry himself, at the time of the ransom demand, admits he believes that the girl is already dead. Does not this possibility or likelihood that the girl is dead destroy the justification for Harry's dirty act? Although it surely would if Harry knew for certain that the girl was dead, I do not think it does insofar as even a small probability of her being saved exists. The reason is that the good to be achieved is so unquestionably good and so passionately felt that even a small possibility of its achievement demands that it be tried. . . .

Once we have satisfied ourselves that a Dirty Harry problem is conceptually possible and that, in fact, we can specify one set of concrete circumstances in which it exists, one might think that the most difficult question of all is, What ought to be done? I do not think it is. I suspect that there are very few people who would not want Harry to do something dirty in the situation specified. I know I would want him to do what he did, and what is more, I would want anyone who policed for me to be prepared to do so as well. Put differently, I want to have as police officers men and women of moral courage and sensitivity.

But to those who would want exactly that, the Dirty Harry problem poses its most irksome conclusion. Namely, that one cannot, at least in the specific case at hand, have a policeman who is both just and innocent. The troublesome issue in the Dirty Harry problem is not whether under some utilitarian calculus a right choice can be made, but that the choice must always be between at least two wrongs. And in choosing to do either wrong, the policeman inevitably taints or tarnishes himself. . . .

THE DIRTY HARRY PROBLEM II: DIRTY MEN AND DIRTY WORK

Dirty Harry problems arise quite often. For policemen, real, everyday policemen, Dirty Harry problems are part of their job and thus considerably more than rare or artificial dramatic exceptions. To make this point, I will translate some rather familiar

police practices, street stops and searches and victim and witness interrogation, into Dirty Harry problems.

Good Ends and Dirty Means

The first question our analysis of street stops and searches and victim and witness interrogation must satisfy is, For policemen, do these activities present the cognitive opportunity for the juxtaposition of good ends and dirty means to their achievement? Although the "goodness" question will be considered in some detail later, suffice it to say here that police find the prevention of crime and the punishment of wrongful or criminal behavior a good thing to achieve. Likewise, they, perhaps more than any other group in society, are intimately aware of the varieties of dirty means available for the achievement of those good ends. In the case of street stops and searches, these dirty alternatives range from falsifying probable cause for a leading stop, to manufacturing a false arrest to legitimate an illegal search, to simply searching without the fraudulent covering devices of either. In the case of victim or witness interrogations, dirty means range all from dramaturgically "chilling" a *Miranda* warning by an edited or unemphatic reading to Harry's grinding a man's bullet-shattered leg to extort a confession from him.

. . . Are not the police . . . less sensitive to such things than are most of us? I think there is no question that they are, and our contention about the prevalence of Dirty Harry problems in policing suggests that they are likely to be. How does this "tough-minded" attitude toward dirty means affect our argument? At least at this stage it seems to strengthen it. That is, the failure of police to regard dirty means with the same hesitation that most citizens do seems to suggest that they juxtapose them to the achievement of good ends more quickly and more readily than most of us. . . .

1. ***The Operative Assumption of Guilt.*** In street stops and searches as well as interrogations, it is in the nature of the police task that guilt is assumed as a working premise. That is, in order for a policeman to do his job, he must, unless he clearly knows otherwise, assume that the person he sees is guilty and the behavior he is witnessing is evidence of some concealed or hidden offense. If a driver looks at him "too long" or not at all or if a witness or suspect talks too little or too much, it is only his operative assumption of guilt that makes those actions meaningful. Moreover, the policeman is often not in a position to suspend his working assumption until he has taken action, sometimes dirty action, to disconfirm it.

2. ***The Worst of All Possible Guilt.*** The matter of the operative assumption of guilt is complicated further because the policeman is obliged to make a still higher-order assumption of guilt, namely, that the person is not only guilty, but dangerously so. In the case of street stops and searches, for instance, although the probability of coming upon a dangerous felon is extremely low, policemen quite reasonably take the possibility of doing so as a working assumption on the understandable premise that once is enough. Likewise the premise that the one who has the most to hide will try hardest to hide it is a reasonable assumption for interrogation.

3. ***The Great Guilty Place Assumption.*** The frequency with which policemen confront the worst of people, places, and occasions creates an epistemological problem of serious psychological proportions. As a consequence of his job, the policeman is constantly exposed to highly selective samples of his environment. That he comes to read a clump of bushes as a place to hide, a roadside rest as a homosexual "tearoom," a sweet old lady as a robbery looking for a place to happen, or a poor young black as someone willing to oblige her is not a question of a perverse, pessimistic, or racist personality, but of a person whose job requires that he strive to see race, age, sex, and even nature in an ecology of guilt. . . .

4. ***The Not Guilty (This Time) Assumption.*** With considerable sociological research and conventional wisdom to support him, the policeman knows that most people in the great guilty place in which he works have committed numerous crimes for which they have never been caught. Thus when a stop proves unwarranted, a search comes up "dry," or an interrogation fails, despite the dirty means, the policeman is not at all obliged to conclude that the person victimized by them is innocent, only that, and even this need not always be conceded, he is innocent this time.

Dirty Means as Ends in Themselves

How do these features of police work, all of which seem to incline police to accept a standard of a guilty victim for their dirty means, bear upon the Dirty Harry problem from which they derive? The most dangerous reading suggests that if police are inclined, and often quite rightly inclined, to believe they are dealing with factually, if not legally, guilty subjects, they become likely to see their dirty acts, not as means to the achievement of good ends, but as ends in themselves—as punishment of guilty people whom the police believe deserve to be punished.

If this line of argument is true, it has the effect, in terms of police perceptions, of moving Dirty Harry problems completely outside of the four-fold table of means–ends combinations created in order to define it. Importantly as well, in terms of our perceptions, Dirty Harry problems of this type can no longer be read as cases of dirty means employed to the achievement of good ends. For unless we are willing to admit that in a democratic society a police arrogates to itself the task of punishing those who they think are guilty, we are forced to conclude that Dirty Harry problems represent cases of employing dirty means to dirty ends, in which case, nobody, not the police and certainly not us, is left with any kind of moral dilemma. . . .

If . . . police collapse the dirty means–good ends division on some occasions and become punishers, . . . on many other occasions, collapse does not occur and Dirty Harry problems, as defined, are still widely possible. What must be remembered next, on the way to making their possibility real, is that policemen know, or think they know, before they employ a dirty means that a dirty means and only a dirty means will work.

Only a Dirty Means Will Work

The moral standard that a policeman know in advance of resorting to a dirty means that a dirty means and only a dirty means will work, rests heavily on two technical dimensions: (1) the professional competence of the policeman and (2) the range of legitimate working options available to him. Both are intimately connected, though the distinction to be preserved between them is that the first is a matter of the policeman's individual competence and the second of the competence of the institutions for which (his department) and with which (the law) the policeman works.

In any concrete case, the relations between these moral and technical dimensions of the Dirty Harry problem are extremely complicated. But a priori it follows that the more competent a policeman is at the use of legal means, the less he will be obliged to resort to dirty alternatives. Likewise, the department that trains its policemen well and supplies them with the resources—knowledge and material—to do their work will find that the policemen who work for them will not resort to dirty means "unnecessarily," meaning only those occasions when an acceptable means will work as well as a dirty one. . . .

Nevertheless . . . Dirty Harry problems can arise wherever restrictions are placed on police methods and are particularly likely to do so when police themselves perceive that those restrictions are undesirable, unreasonable, or unfair. . . .

Compelling and Unquestionable Ends

If Dirty Harry problems can be shown to exist in their technical dimensions—as genuine means–ends problems where only dirty means will work—the question of the magnitude and urgency of the ends that the dirty means may be employed to achieve must still be confronted. Specifically, it must be shown that the ends of dirty means are so desirable that the failure to achieve them would cast the person who is in a position to do so in moral disrepute.

The two most widely acknowledged ends of policing are peace keeping and law enforcement. It would follow, of course, that if both these ends were held to be unworthy, Dirty Harry problems would disappear. There are arguments challenging both ends. For instance, certain radical critiques of policing attempt to reduce the peace-keeping and law-enforcing functions of the police in the United States to nothing more than acts of capitalist oppression. From such a position flows not only the denial of the legitimacy of any talk of Dirty Harry problems, but also the denial of the legitimacy of the entire police function. . . .

If we can conclude that peace and punishment are ends unquestionable and compelling enough to satisfy the demands of Dirty Harry problems, we are led to one final question on which we may draw from some sociological theories of the police for assistance. If the Dirty Harry problem is at the core of the police role, or at least near to it, how is it that police can or do come to reconcile their use of—or their failure to use—dirty means to achieve unquestionably good and compelling ends?

PUBLIC POLICY AND POLICE MORALITY: THREE DEFECTIVE RESOLUTIONS OF THE DIRTY HARRY PROBLEM

The contemporary literature on policing appears to contain three quite different types of solution or resolution. But because the Dirty Harry problem is a genuine moral dilemma, that is, a situation which will admit no real solution or resolution, each is necessarily defective. Also, understandably, each solution or resolution presents itself as an answer to a somewhat different problem. In matters of public policy, such concealments are often necessary and probably wise . . . In discovering that each is flawed and in disclosing the concealments which allow the appearance of resolution, . . . we only wish to make clear what is risked by each concealment and to face candidly the inevitably unfortunate ramifications which must proceed from it.

The Professional Model

In the works of August Vollmer, Bruce Smith, O. W. Wilson, and those progressive police administrators who still follow their lead, a vision of the perfect police agency and the perfect policeman has gained considerable ground. Labeled "the professional model" in police circles—though entirely different from any classical sense of profession or professional—it envisions a highly trained, technologically sophisticated police department operating free from political interference with a corps of well-educated police responding obediently to the policies, orders, and directives of a central administrative command. It is a vision of police officers . . . as "snappy bureaucrats," cogs in a quasi-military machine who do what they are told out of a mix of fear, loyalty, routine, and detailed specification of duties.

The professional model, unlike other solutions to be considered, is based on the assumption that the policeman's motives for working can be made to locate within his department. He will, if told, work vice or traffic, juvenile or homicide, patrol passively or aggressively, and produce one, two, four, or six arrests, pedestrian stops, or reports per hour, day, or week as his department sees fit. In this way the assumption and vision of the professional model in policing is little different from that of any bureaucracy which seeks by specifying tasks and setting expectations for levels of production—work quotas—to coordinate a regular, predictable, and efficient service for its clientele.

The problem with this vision . . . is that when the product to be delivered is some form of human service—education, welfare, health, and police bureaucracies are similar in this way—the vision seems always to fall short of expectations. On the one hand the would-be bureaucratic operatives—teachers, social workers, nurses, and policemen—resent being treated as mere bureaucrats and resist the translation of their work into quotas, directives, rules, regulations, or other abstract specifications. On the other hand, to the extent that the vision of an efficient and obedient human service bureaucracy is realized, the clientele of such institutions typically come away with the impression that no one in the institution truly cares about their problems. And, of course, in that the aim of bureaucratization is to locate employees' motives for work within the bureaucracy, they are absolutely correct in their feelings.

To the extent that the professional model succeeds in making the ends of policing locate within the agency as opposed to moral demands of the tasks which policemen are asked by their clients to do, it appears to solve the Dirty Harry problem. When it succeeds, it does so by replacing the morally compelling ends of punishment and peace with the less human, though by no means uncompelling, ends of bureaucratic performance. However, this resolution certainly does not imply that dirty means will disappear, only that the motives for their use will be career advancement and promotion. Likewise, on those occasions when

a morally sensitive policeman would be compelled by the demands of the situational exigencies before him to use a dirty means, the bureaucratic operative envisioned by the professional model will merely do his job. Ambitious bureaucrats and obedient timeservers fail at being the type of morally sensitive souls we want to be policemen. The professional model's bureaucratic resolution of the Dirty Harry problem fails in policing for the same reason it fails in every other human service agency; it is quite simply an impossibility to create a bureaucrat who cares for anything but his bureaucracy.

The idealized image of the professional model, which has been responded to with an ideal critique, is probably unrealizable. Reality intervenes as the ideal type is approached. The bureaucracy seems to take on weight as it approaches the pole, is slowed, and may even collapse in approaching.

Peace, not Punishment, as Goal

A second effort in the literature of contemporary policing also attempts to address the Dirty Harry problem by substituting an alternative to the presently prevailing police ends of punishment. Where the professional model sought to substitute bureaucratic rewards and sanctions for the moral end of punishment, . . . Egon Bittner . . . seek[s] to substitute the end of peace. In beautifully chosen words, examples, and phrasing, Bittner leads his readers to conclude that peace is historically, empirically, intellectually, and morally the most compelling, unquestionable, and humane end of policing. Bittner is, I fear, absolutely right.

It is the end of peace which legitimates the extension of police responsibilities into a wide variety of civil matters—neighborhood disputes, loud parties, corner lounging, lovers' quarrels, political rallies, disobedient children, bicycle registration, pet control, and a hundred other types of tasks which a modern "service" style police department regularly is called upon to perform. With these responsibilities, which most "good" police agencies now accept willingly and officially, also comes the need for an extension of police powers. Arrest is, after all, too crude a tool to be used in all the various situations in which our peace-keeping policemen are routinely asked to be of help. "Why should," asks Herman Goldstein, in a manner in which Bittner would approve, "a police officer arrest and charge a disorderly tavern patron if ordering him to leave the tavern will suffice? Must he arrest and charge one of the parties in a lovers' quarrel if assistance in forcing a separation is all that is desired?" There is no question that both those situations could be handled more peacefully if police were granted new powers which would allow them to handle those situations in the way Goldstein rhetorically asks if they should. That such extensions of police powers will be asked for by our most enlightened police departments in the interests of keeping the peace is absolutely certain. If the success of the decriminalization of police arrests for public intoxication, vagrancy, mental illness, and the virtually unrestricted two-hour right of detention made possible by the Uniform Law of Arrest are any indication of the likelihood of extensions being received favorably, the end of peace and its superiority over punishment in legitimating the extension of police powers seem exceedingly likely to prevail further.

The problem with peace is that it is not the only end of policing so compelling, unquestionable, and in the end, humane. Amid the good work toward the end of peace that we increasingly want our police to do, it is certain that individuals or groups will arise who the police, in all their peace-keeping benevolence, will conclude, on moral if not political or institutional grounds, have "got it coming." And all the once dirty means which were bleached in the brilliant light of peace will return to their true colors.

The Police Officer as Craftsman

The third and final attempt to resolve the Dirty Harry problem is offered by Jerome Skolnick, who in *Justice Without Trial* comes extremely close to stating the Dirty Harry problem openly when he writes:

> . . . He (the policeman) sees himself as a craftsman, at his best, a master of his trade . . . [he] draws a moral distinction between criminal law and criminal procedure. The distinction is drawn somewhat as follows: The substantive law of crimes is intended to control the behavior of people who willfully injure

persons or property, or who engage in behaviors having such a consequence, such as the use of narcotics. Criminal procedure, by contrast, is intended to control authorities, not criminals. As such, it does not fall into the same *moral* class of constraint as substantive criminal law. If a policeman were himself to use narcotics, or to steal, or to assault, *outside the line of duty,* much the same standards would be applied to him by other policemen as to the ordinary citizen. When, however, the issue concerns the policeman's freedom to carry out his *duties,* another moral realm is entered.[2]

What is more, Skolnick's craftsman finds support from his peers, department, his community, and the law for the moral rightness of his calling. He cares about his work and finds it just.

What troubles Skolnick about his craftsman is his craft. The craftsman refuses to see, as Skolnick thinks he ought to, that the dirty means he sometimes uses to achieve his good ends stand in the same moral class of wrongs as those he is employed to fight. Skolnick's craftsman reaches this conclusion by understanding that his unquestionably good and compelling ends, on certain occasions, justify his employment of dirty means to their achievement. Skolnick's craftsman, as Skolnick understands him, resolves the Dirty Harry problem by denying the dirtiness of his means.

. . . The problem is not that the craftsman will sometimes choose to use dirty means. If he is morally sensitive to its demands, every policeman's work will sometimes requires as much. What is worrisome about Skolnick's craftsman is that he does not regard his means as dirty and, as Skolnick tells us, does not suffer from their use. The craftsman, if Skolnick's portrait of him is correct, will resort to dirty means too readily and too easily. He lacks the restraint that can come only from struggling to justify them and from taking seriously the hazards involved. . . .

The alternative the Dirty Harry problem leads us to is ensuring that the craftsman regards his dirty means as dirty by applying the same retributive principles of punishment to his wrongful acts that he is quite willing to apply to others! It is, in fact, only when his wrongful acts are punished that he will come to see them as wrongful and will appreciate the genuine moral—rather than technical or occupational—choice he makes in resorting to them. The prospects for punishment of such acts are by no means dim, and considerable strides in this area have been made. It requires far fewer resources to punish than to reward. Secondly, the likelihood that juries in civil suits will find dirty means dirtier than police do is confirmed by police claims that outsiders cannot appreciate the same moral and technical distinctions that they do. Finally, severe financial losses to police agencies as well as to their officers eventually communicate to both that vigorously policing themselves is cheaper and more pleasing than having to pay so heavily if they do not. If under such conditions our craftsman police officer is still willing to risk the employment of dirty means to achieve what he understands to be unquestionably good ends, he will not only know that he has behaved justly, but that in doing so he must run the risk of becoming genuinely guilty as well.

A FINAL NOTE

In urging the punishment of policemen who resort to dirty means to achieve some unquestionably good and morally compelling end, we recognize that we create a Dirty Harry problem for ourselves and for those we urge to effect such punishments. It is a fitting end, one which teaches once again that the danger in Dirty Harry problems is never in their resolution, but in thinking that one has found a resolution with which one can truly live in peace.

NOTES

1. "Dirty" here means both "repugnant" in that it offends widely shared standards of human decency and dignity and "dangerous" in that it breaks commonly shared and supported norms, rules, or laws for conduct.

2. Jerome Skolnick, *Justice Without Trial,* 2nd ed. (New York: John Wiley & Sons, 1975), p. 182.

REVIEW AND DISCUSSION QUESTIONS

1. What did Dirty Harry do in the movie of that name? What, then, is a Dirty Harry problem? Do you agree that police officers often face Dirty Harry problems?
2. What three questions must be answered (or what three things must be known) for a Dirty Harry problem to exist?
3. The author believes that Harry did what he should have done (and that almost everybody would agree with Harry's using dirty means). Do you agree? If so, why is there a Dirty Harry problem?
4. Klockars argues that if the police believe that they are dealing with guilty people, then this fact may imply that an apparent Dirty Harry problem is no longer a genuine moral dilemma. Explain.
5. Explain each of the three responses that Klockars discusses to the Dirty Harry problem. What shortcomings, if any, do you see in each of them?
6. Can the Dirty Harry Problem be solved? What, if anything, can police departments, the law, or society in general do to address problem? Can better police training and increased resources for police departments eliminate the problem? How severely should police officers be punished for using dirty means to promote good ends?

A Debate on the Exclusionary Rule

Malcolm Richard Wilkey and Stephen H. Sachs

The Fourth Amendment protects us from unreasonable searches and seizures, and it requires that police warrants be issued only for probable cause. The important case of *Mapp v. Ohio* provides a good example of an unconstitutional search and seizure. Here is how the Supreme Court described the events leading to the arrest of a certain Miss Mapp for having obscene material in her possession.

> On May 23, 1957, three Cleveland police officers arrived at appellant's residence in that city pursuant to information that "a person [was] hiding out in the home, who was wanted for questioning in connection with a recent bombing, and that there was a large amount of policy paraphernalia being hidden in the home." Miss Mapp and her daughter by a former marriage lived on the top floor of the two-family dwelling. Upon their arrival at that house, the officers knocked on the door and demanded entrance but appellant, after telephoning her attorney, refused to admit them without a search warrant. They advised their headquarters of the situation and undertook a surveillance of the house.
>
> The officers again sought entrance some three hours later when four or more additional officers arrived on the scene. When Miss Mapp did not come to the door immediately, at least one of the several doors to the house was forcibly opened and the policemen gained admittance. Meanwhile Miss Mapp's attorney arrived, but the officers, having secured their own entry, and continuing in their defiance of the law, would permit him neither to see Miss Mapp nor to enter the house. It appears that Miss Mapp was halfway down the stairs from the upper floor to the front door when the officers, in this highhanded manner, broke into the hall. She demanded to see the search warrant. A paper, claimed to be a warrant, was held up by one of the officers. She grabbed the "warrant" and placed it in her bosom. A struggle ensued in which the officers recovered the piece of paper and as a result of which they handcuffed appellant because she had been "belligerent" in resisting their official rescue of the "warrant" from her person. Running roughshod over appellant, a policeman "grabbed" her, "twisted [her] hand," and she "yelled [and] pleaded with him" because "it was hurting." Appellant, in handcuffs, was then forcibly taken upstairs to her bedroom where the officers searched a dresser, a chest of drawers, a closet and some suitcases. They also looked into a photo album and through personal papers belonging to the appellant. The search spread to the rest

of the second floor including the child's bedroom, the living room, the kitchen and a dinette. The basement of the building and a trunk found therein were also searched. The obscene materials for possession of which she was ultimately convicted were discovered in the course of that widespread search.[1]

To uphold the Fourth Amendment and deter the police from violating it, the Supreme Court has held in *Mapp* and other cases that illegally obtained evidence should be excluded from criminal trial; that is, it should not be heard by the jury. This is called the exclusionary rule, and violation of it is grounds for an appellate court to overturn a conviction. Sometimes, of course, the excluded evidence may not only be reliable but may also be necessary to convict an undeniably guilty defendant such as Miss Mapp. Preventing the jury from hearing such evidence results in a guilty person going free. For this and other reasons, Malcolm Richard Wilkey, an appellate court judge, argues that the rule is irrational and should be replaced by a system that enforces the Fourth Amendment without letting guilty people off. Stephen H. Sachs, a Maryland prosecutor, contends, to the contrary, that the rule results in relatively few criminals being set free and that it is the only effective means of deterring the police and prosecutors from violating our constitutional rights.

Wilkey: A Critique of the Exclusionary Rule

WHAT I PROPOSE

The time has come to abolish the exclusionary rule. This judicial creation has failed. Its alleged value has always hinged on the assumption, still unproven after seventy years, that it will deter police illegality. This doubtful gain is overwhelmed by the rule's horrible side effects, which are so destructive of the efficiency and integrity of our criminal justice system and whose costs are undeniable. Fourth Amendment suppression motions arise so frequently that, whether granted or not, they constitute the single issue most burdensome to the courts in criminal trials. And when a motion is granted the result is a manifest injustice—the exclusion of reliable evidence which may be essential to convicting the undeniably guilty. Moreover, the exclusionary rule perversely insures that the wrongdoer, the police officer who conducted the illegal search, escapes punishment as well. Under the rule's logic, two wrongs apparently make a right.

Abolition of the rule does not mean abolition of our constitutional rights. The exclusionary rule is not the Fourth Amendment, and it is not a constitutional necessity. It is clear from the way the rule originated in *Weeks v. United States* in 1914 that the Court was simply choosing a method of enforcing the Fourth Amendment. The search and seizure opinions of the Supreme Court for the last twenty years exhibit a reiteration by one Justice after another that this is *a* method of enforcing a constitutional protection, not necessarily *the* precise method commanded by the Constitution but instead one chosen by the High Court. . . . The Constitution prohibits unreasonable searches and seizures: It may mandate *a* remedy to enforce that prohibition, but nowhere does the Constitution mandate the exclusion of all other possible remedies.

Three workable legislative alternatives are available: an independent review board in the executive branch; a civil tort action against the government; and a mini-trial separate from the main criminal trial. One alternative which should *not* be adopted is the "good faith" exception advocated by the Attorney General's Task Force on Violent Crime. Although appealing to the extent that it reduces the number of guilty criminals who go free, the good faith proposal perpetuates all the other evils of the exclusionary rule and adds new additional costs.

Of all the civilized countries in the world, only the United States applies this demonstrably illogical exclusionary remedy to illegal searches and seizures of material evidence. The Congress should act decisively to repudiate the absurd notion that the United States alone is incapable of using its executive, administrative, and judicial processes *directly* to control the conduct of its law enforcement officials.

COSTS OF THE PRESENT IRRATIONALITY

In my monograph on the Fourth Amendment I have described at some length twelve costs of the exclusionary rule. Without going into a detailed description of each, well known to students of the exclusionary rule, let me enumerate these twelve undeniable costs for reference:

> Cost 1: "The criminal is to go free because the constable has blundered."
>
> Cost 2: Only the undeniably guilty benefit from the exclusionary rule, while innocent victims of illegal searches have neither protection nor remedy.
>
> Cost 3: The exclusionary rule in any form vitiates all internal disciplinary efforts by law enforcement agencies.
>
> Cost 4: The disposition of exclusionary rule issues constitutes an unnecessary and intolerable burden on the court system.
>
> Cost 5: The exclusionary rule forces the judiciary to perform the executive branch's job of disciplining its employees.
>
> Cost 6: The misplaced burden on the judiciary deprives innocent defendants of due process.
>
> Cost 7: The exclusionary rule encourages perjury by the police.
>
> Cost 8: The exclusionary remedy makes hypocrites out of judges.
>
> Cost 9: The high cost of applying the exclusionary rule causes the courts to expand the scope of search and seizure for all citizens.
>
> Cost 10: The exclusionary remedy is applied with no sense of proportion to the crime of the accused.
>
> Cost 11: The exclusionary remedy is applied with no sense of proportion to the misconduct of the officer.
>
> Cost 12: All of the above costs result inevitably in greatly diminished respect for the judicial process among lawyers and laymen alike.[2]

The astounding fact about all twelve costs is that they are undenied and undeniable. . . . The proponents of preserving the exclusionary remedy are forced to admit the existence of these high costs in our criminal justice system, yet they blithely assume that the only purported rationale of the rule outweighs these admitted costs.

IMPACT ON THE COURTS

While the most dramatic cost of the exclusionary remedy has been that "the criminal is to go free because the constable has blundered," this should not overshadow the tremendous impact that the rule has on the sheer work load of the courts. A recent General Accounting Office (GAO) study, though focusing on the intake of cases in selected United States Attorneys' offices, reveals the enormous burden that the exclusionary rule has placed on the trial and appellate courts. According to the GAO data, Fourth Amendment suppression motions are far and away the most common issue in the trial of criminal cases. As was stated on the first page of the GAO report, "*thirty-three percent* of the defendants who went to trial filed Fourth Amendment suppression motions."...

The Fourth Amendment suppression motion is the most important and frequent single issue arising in criminal trials. The GAO study illustrates this: 60.1 percent of all motions filed involved the Fourth Amendment. The next most numerous type of motion—dealing with confessions—amounted to only 23.2 percent: hence there is no single legal issue which even comes close to Fourth Amendment search and seizure motions for importance in the trial of criminal cases. Irrespective of whether the evidence is excluded or not, the burden on the trial court is undeniable. . . .

RATIONALE FOR IRRATIONALITY

The sole rationale for preserving the rule is now, and has been at least since *Linkletter v. Walker,* that it deters police misconduct. Efforts to resurrect other rationales, long since discarded by the Supreme Court, are a very revealing admission by even the rule's most ardent supporters that deterrence has *not* been proved by seventy years of experience, and such efforts arouse a sneaking suspicion that the supporters know deterrence is impossible to prove. . . .

[A]s Justice Blackmun remarked in *United States v. Janis* in 1976, "No empirical researcher, proponent or opponent of the rule, has yet been able to establish with any assurance whether the rule has a deterrent effect even in situations in which it is now applied." On this same point,

Chief Justice Burger in *Bivens* described the current situation:

> I do not question the need for some remedy to give meaning and teeth to the constitutional guarantees against unlawful conduct by government officials. . . . But the hope that this objective could be accomplished by the exclusion of reliable evidence from criminal trials was hardly more than a wistful dream. . . . There is no empirical evidence to support the claim that the rule actually deters illegal conduct of law enforcement officials. . . . We should view the suppression doctrine as one of the experimental steps in the great tradition of the common law and acknowledge its shortcomings. But in the same spirit we should be prepared to discontinue what the experience of over half a century has shown neither deters errant officers nor affords a remedy to the totally innocent victim of official misconduct.

It is highly significant that, in nearly seventy years of Supreme Court decisions on the exclusionary rule, the High Court itself has never cited empirical data to support its faith that the rule actually deters. Indeed, the fragmented nature itself of law enforcement and the pressures generated by the exclusionary rule may make it logically impossible for it to have any appreciable deterrent effect.

And even if there were shown to be *some* deterrent effect, supporters of the rule have never addressed what would seem to be a primary question: How much deterrence is necessary to outweigh the admitted cost of the rule? . . .

BURDEN OF PROOF

. . . Nowhere else in our system of proof do we bar valid, probative, irrefutable, relevant, material evidence on some unproven nebulous theory of overall gain in the system. Those who want to bar this evidence never seem to realize that the *burden of proving* that some greater good will be gained is on them.

Evidentiary privileges, it must be conceded, constitute another exception to the general rule of law that valid probative evidence of unquestioned truth is to be admitted at trial. However, a careful analysis of these privileges reveals that they are grounded in a rationale which cannot be used to sustain the exclusionary rule.

The overriding policy underlying the basic evidentiary privileges is the preservation of confidentiality. Many of the privileges exist to preserve confidential interpersonal relationships—for example, husband-wife, lawyer-client, doctor-patient. The other privileges are aimed at preventing the release of information, for various and obvious policy reasons. Thus we have privileges for the identity of a police informer, for the political votes cast by a person, for trade secrets, and for state secrets and other official information.

In each of these cases, the legal system makes the rational evaluation that the preservation of confidentiality is of greater public interest than the availability of the evidence thereby excluded. The crucial fact about making this determination in any of these situations is that *there is no alternative.* If the evidence or testimony is required in open court, the underlying confidentiality interest vanishes. Thus there is a stark choice between disclosure or a privilege. Whatever the decision, there will be a gain and a cost, and we can hope only that we can achieve a net benefit.

The case of the exclusionary rule is exactly the opposite. If the evidence is admitted, the Fourth Amendment interest at stake is not inherently and irretrievably forsaken. *Alternatives do exist.* The police officer may be disciplined, and the victim of illegal search may recover damages. Moreover, when we recognize an evidentiary privilege, the benefit gained is clear and tangible: The confidential information is protected. But with the exclusionary rule we are asked to put our faith behind the assumption that if we throw out enough solid evidence and let enough guilty criminals go free, we ultimately will achieve a greater protection of Fourth Amendment rights. The dubious nature of this reasoning, when combined with a recognition of the available alternatives for *directly* vindicating the Fourth Amendment interests at stake, makes clear that we are throwing away important and irrefutable evidence for no good reason. . . .

NO RESTORATION OF PRE-SEARCH SITUATION

One of the most astonishing assertions by the contemporary supporters of the exclusionary rule is that the remedy simply restores the situation between the government and the target of the

illegal search to precisely what it was before the violation of the Fourth Amendment took place, that is, that matters are returned to how they would have been had the constable not blundered. This assertion is simply divorced from the reality of what goes on in a Fourth Amendment search and seizure case. The exclusionary rule never restores the parties to where they were before the search.

In at least 75 to 80 percent of the search and seizure cases coming into the federal court, the objects of the search are contraband items: narcotics, smuggled goods, illegal weapons, gambling paraphernalia, or stolen property. These are *never* returned to the possessor from whom they were taken at the time of the illegal search, so this party is hardly in the same situation as before. On the other hand not only does the government lose the prosecution based on the unquestioned material evidence which is secured, the government may also be precluded by the doctrine of "the fruit of the poisonous tree" from prosecuting a defendant on the basis of other evidence which it could have obtained validly. So the exclusionary rule never restores the prior status of either party.

The fact that the situation is *not* restored to its pristine condition points up another illogicality of the exclusionary rule. If the government can take away an unlicensed pistol from a person on the street and confiscate it, and yet not strike a damaging blow to the integrity of justice, why should it be demeaning and unworthy of the government to prosecute the man who carried the pistol? If the government seizes, legally or illegally, a million dollars worth of cocaine, it has the right and statutory duty to destroy that cocaine, not to restore it to its original owner. If this is not an affront to our concept of justice, why would it be an affront of the judicial system to prosecute the smuggler of the cocaine? . . .

ABSENCE OF THE RULE IN OTHER CIVILIZED COUNTRIES

To my mind, one of the most damning indictments of the exclusionary rule has been the fact that, while we have had it for nearly seventy years, no other civilized country has adopted our mandatory rule of the exclusion of probative material evidence because of police misconduct. . . .

This failure to follow our example has not been mere happenstance. Law reform commissions in England, Canada, Israel, and other countries have studied the application and the impact of our exclusionary rule and have not adopted it. . . . These countries do discipline their police, they do protect their citizens from illegal searches and seizures and from police harassment, and they do this by methods other than the irrational exclusion of evidence. The obvious existence of other civilized methods of protecting the citizenry against illegal police action totally refutes the idea that there are no viable alternatives to the exclusionary rule. . . .

ALTERNATIVES

. . . [T]here are several alternative choices for enforcing the Fourth Amendment protection against unreasonable searches and seizures. Each of these I have discussed at some length in my monograph.[3] To enumerate: The first and most logical alternative would be to adopt a system under which the executive branch disciplines itself. While the judiciary may ultimately have the responsibility for implementing all constitutional protections, this does not mean that the judiciary must intervene in every single instance of an alleged violation, as happens now when the exclusionary rule functions as the sole method of enforcing the Fourth Amendment. While individual law enforcement agencies may be too close to their own people to discipline them effectively all the time, this does not preclude the setting up of an overall disciplinary board or agency in the executive branch, where the alleged misdeeds of enforcement officials can be investigated, tried, and punished appropriately without any impact on the trial of the accused for his or her crime.

Second, surely a civil tort remedy can be created, under the Federal Tort Claims Act or elsewhere, which would give victims of searches and seizures a claim remedy against erring police officers. Under the present exclusionary rule, no one is compensated in any way for a constitutional violation unless he or she is actually guilty of possessing incriminating material objects and is successful in having them suppressed. If the search turns up nothing incriminating, no matter how outrageous the violation, the victim of the search has no remedy whatsoever.

The above two remedies, in general outline, are the remedies usually relied upon in other countries in the world with judicial systems similar to ours—in England, Israel, and the British Commonwealth nations—to discipline their law enforcement personnel and to protect the citizenry against illegal searches and seizures. There is no reason why similar remedies could not work in the United States of America. . . .

I have suggested a third method of enforcement which, even in the absence of any congressional legislation on the matter, could be implemented by our judicial system to enforce the Fourth Amendment in a much more logical fashion. There could be a mini-trial of the alleged offending officer after the trial of the accused for his substantive crime. If a motion charging the officer with an illegal search were made, the trial judge would reserve judgment on that extraneous matter until the guilt or innocence of the accused had been established. The evidence seized would come in, for the officer's conduct would be totally irrelevant to the question of whether the narcotics, the unlicensed gun, or the contraband had been found in the possession of the accused, that is, the conduct would be irrelevant to the guilt of the defendant. The material evidence would be admitted, though, only on condition that the conduct of the officer be appraised later. After subsequent judicial inquiry into the officer's conduct, if the officer overstepped legal bounds, the trial judge would then inform the agency that the conviction of the accused would stand only if proportional disciplinary measures were taken against the officer and if a report of these was provided to the court within a given time. But if the agency did not discipline the officer sufficiently to act as a deterrent—not only to his, but his colleagues', future misconduct of this type—then the motion of the defense to exclude the seized evidence would be granted under the present exclusionary rule. The result in the usual case would be insufficient evidence to sustain the conviction. Of course, if the trial judge found that the officer's conduct was well within Fourth Amendment standards, then there would be no exclusion of the evidence and no required punishment of the officer.

This, I submit, would form an effective and rational deterrent to law enforcement officials' violations of the Fourth Amendment. It would result in freeing the criminal only if the law enforcement agency were recalcitrant and refused to punish its erring officer. Assuming that the law enforcement agency took disciplinary action, this method would not free the guilty, but instead would result in the proportionate punishment of the erring officer while the obviously guilty also received his or her just deserts. It is my belief that if a United States district court inaugurated a system of required investigation and punishment of alleged Fourth Amendment violators—separate and distinct from the trial and conviction or acquittal of the accused—the Supreme Court ultimately would sustain this as a far more effective and logical method of enforcing the Fourth Amendment.

Sachs: A Defense of the Exclusionary Rule

Like most law enforcement officials, I share our citizens' alarm and anger at the dreadful impact of crime in our society and want to strengthen law enforcement's capabilities. I agree with many of the recommendations made to Attorney General Smith by his Task Force on Violent Crime, including those which facilitate pretrial detention of some demonstrably dangerous criminals and curb collateral review of state convictions by federal courts. Furthermore, I deplore the hypertechnical excess with which some courts have interpreted the *substance* of the Fourth Amendment's guarantees against unreasonable searches and seizures. . . .

What I also believe, however, is that the exclusionary rule, the *remedy* for a Fourth Amendment violation which suppresses its fruits and denies government the benefit of its unconstitutional conduct, is sound in theory and effective in practice.

The rule is also very fragile, especially in today's atmosphere of understandable public outrage at

crime and at our perceived inability to do much about it. It is vulnerable to attack because its values are abstract while its costs are tangible. It frequently excludes hard evidence—the truth—from trial. It appears to reward the undeserving criminal, whom it sometimes frees because "the constable blundered." It seems to give aid and comfort only to the enemy in the war on crime. It makes almost no sense to citizens fed up with crime and impatient with legal "technicalities" who want to believe that crime would disappear if only courts would stop coddling criminals. That is why the rule, although it has plenty of responsible critics, has become a favorite whipping boy of anticrime rhetoricians.

My purpose in this essay is to bear witness to what my own experience and study have taught me:

1. the rule is of constitutional origin and beyond the reach of Congress;

2. it results in freeing guilty criminals in a relatively small proportion of cases;

3. it definitely deters police and prosecutor violations of constitutional rights to privacy;

4. it manifests our refusal to stoop to conquer, to convict lawbreakers by relying on official lawlessness, a vital demonstration of our commitment to the rule of law. . . .

THE CONSTITUTIONAL STATUS OF THE EXCLUSIONARY RULE

I believe that the exclusionary rule is a rule of constitutional origin which only the Supreme Court, as part of its process of constitutional adjudication, can alter. The rule was first announced in *Weeks v. United States,* where Justice Day focused on the *constitutional necessity* of effectively enforcing the Fourth Amendment:

> The effect of the Fourth Amendment is to put the courts of the United States and Federal officials in the exercise of their power and authority, under limitations and restraints as to the exercise of such power and authority. . . . This protection reaches all alike, whether accused of crime or not, and the duty of giving to it force and effect is obligatory upon all intrusted under our Federal system with the enforcement of the laws. The tendency of those who execute the criminal laws of the country to obtain conviction by means of unlawful seizures and enforced confessions, the latter often obtained after subjecting accused persons to unwarranted practices destructive of rights secured by the Federal Constitution, *should find no sanction in the judgments of the courts,* which are charged at all times with the support of the Constitution, and to which people of all conditions have a right to appeal for the maintenance of such fundamental rights. [Emphasis added by Sachs.]

The Supreme Court emphatically established the exclusionary rule's constitutional roots when it made the rule applicable to the states in *Mapp v. Ohio:*

> We hold that all evidence obtained by searches and seizures in violation of the Constitution is, *by that same authority,* inadmissible in a state court. . . . Moreover, out holding that *the exclusionary rule is an essential part of both the Fourth and Fourteenth Amendments* is not only the logical dictate of prior cases, but it also makes very good sense. [Emphasis added by Sachs.]. . .

While the Court has occasionally limited the reach of the rule, it has not retreated from its constitutional base. The rule may be a "judicially created remedy," but it has clear constitutional underpinnings.

EXAGGERATED CLAIMS ABOUT THE CONSEQUENCES OF THE EXCLUSIONARY RULE

The most severe indictment of the rule and, in the present climate of fear about crime the most damaging one, is that it releases hordes of dangerous and demonstrably guilty criminals to prey upon society.

I believe that such claims are greatly exaggerated. In cases where unconstitutionally seized evidence is indispensable to conviction, convictions will of course be lost. Police and prosecutors will be frustrated. In the occasional celebrated case, the public will be angry. But a recent study by the Comptroller General of the United States proves that the rule operates to free federal criminal suspects in only a tiny percentage of cases. The Comptroller General studied the rule's operation in 2,804 cases in 38 representative United States

Attorneys Offices throughout the country during the period July 1 through August 31, 1978. Of all the cases presented to those federal prosecutors for prosecution, *only 0.4 percent* were declined by the prosecutors because of Fourth Amendment search and seizure problems. Evidence was excluded at trial as a result of Fourth Amendment motions in *only 1.3 percent* of the cases. And over 50 percent of the few defendants whose suppression motions were granted in whole or in part were nonetheless convicted. This data is powerful proof that at least in federal prosecutions—the only cases, it should be remembered, to which proposed federal legislation would apply—the specter of muggers, rapists, and dope pushers set loose by the rule is the product of rhetoric that is not in touch with reality.

To my knowledge, no comparable study of state and local experience has been conducted. Based on my personal observation and my contacts with other Maryland prosecutors, however, I would expect the impact of the rule on state prosecutions to be somewhat greater, but still fall far short of justifying the claims of the alarmists. I believe a fair summary of the Maryland experience to be that the rule has small effect on either declinations by prosecutors or loss at trial in almost all categories of crime, including violent crime, although it should be added that a higher percentage of suppression motions is granted in drug possession cases resulting from spontaneous street encounters than in other types of cases. As Professor LaFave has written in his treatise on search and seizure: "There is reason to believe that the 'cost' of the exclusionary rule, in terms of acquittals or dismissed cases, is much lower than is commonly assumed."[4]

There is an even more fundamental response to the claim that the exclusionary rule frees criminals. As pointed out by Yale Kamisar, we tend to examine the legality of a search only after it has occurred—and after incriminating evidence has been discovered—and therefore lose sight of the fact that if the police had obeyed the Fourth Amendment they would not have the evidence at all. Thus, it is the guarantee of the Fourth Amendment, a cornerstone of our personal liberty, which constricts the police. The rule serves only to put the police in the same position they would have been in had they done what the Constitution commanded. As Kamisar asks: "If the government could not have gained a conviction had it obeyed the Constitution, why should it be permitted to prevail because it has violated the Constitution?"[5]

AN EFFECTIVE DETERRENT

A fundamental assumption underlying creation of the rule was the judicial belief that it deters official lawlessness and "compel[s] respect for the constitutional guaranty in the only effectively available way—by removing the incentive to disregard it." Critics charge that the rule fails in this primary purpose and point to various empirical studies which, they argue, demonstrate that the rule is ineffective as a deterrent. The studies have been thoroughly dissected elsewhere. I have read most of them and agree with Justice Powell's conclusion that they are "inconclusive."

I can't offer statistical studies on the deterrent effect of the rule. What I can offer, however, is my testimony that I have watched the rule deter, routinely, throughout my years as a prosecutor. When an Assistant United States Attorney, for example, advises an FBI agent that he lacks probable cause to search for bank loot in a parked automobile unless he gets a better "make" on the car; or that the agent has a "staleness" problem with the probable cause to believe that the ski masks used in the robbery are still in the suspect's friend's apartment; or that he should apply for a search warrant from a magistrate and not rely on the "consent" of the suspect's sister to search his home—the rule is working. The principal, perhaps the only, reason those conversations occur is that the assistant and the agent want the search to stand up in court.

Episodes like these are commonplace. They are part of the routine of every federal prosecutor with whom I have worked. Although my present office has more limited criminal jurisdiction, such police-prosecutor consultation is customary in all of our cases when Fourth Amendment concerns arise. I strongly suspect that scenes like these are repeated daily throughout federal law enforcement and on homicide, narcotics, and gambling squads in cities throughout the country. In at least three Maryland jurisdictions, for example, prosecutors are on twenty-four hour call to field search and seizure questions presented by police officers.

These contracts do not occur because of some self-limiting controls in the police and prosecutors themselves. I hope and trust that most of us in law enforcement are principled enough to avoid violating the clear constitutional rights of suspects. But in the heat of the chase, and in the absence of effective sanction, I believe that we would define those rights somewhat narrowly. Questions of adequate identification, "staleness" of information, and the need for a warrant will be answered differently by unchecked law enforcers than by judges. We are, after all, hunters stalking crime. It is simply too much to ask for objectivity in the midst of the hunt, especially when the quarry is in sight. This is precisely what the warrant requirement of the Fourth Amendment is about. As Justice Jackson once put it:

> The point of the Fourth Amendment, which often is not grasped by zealous officers, is not that it denies law enforcement support of the usual inferences which reasonable men draw from evidence. Its protection consists in requiring that those inferences be drawn by a neutral and detached magistrate instead of being judged by the officer engaged in the often competitive enterprise of ferreting out crime.

Exclusion from evidence is almost certainly the only effective deterrent in the vast majority of unconstitutional intrusions. Even critics of the rule are quick to acknowledge the severe limitations of police self-discipline or court damage actions as deterrents when crime-fighting police officers are in the dock. In rare cases involving especially gross misconduct, a police disciplinary board or a court or jury in a damage action might impose sanctions, at least if the victim of the trespass is innocent and the police misconduct truly outrageous.

But most of the suppression cases do not deal with such outrageous conduct. They deal with undramatic Fourth Amendment concerns—the sufficiency of "probable cause" in a given case, whether "exigent circumstances" excuse the necessity of a warrant, whether there is sufficient corroboration of the tip of an anonymous informer to justify intrusion into a suspect's apartment. These requirements are not the stuff to move police disciplinary boards, or judges and juries accustomed to awarding damages on the basis of "fault." But they are our constitutional rules of the road and only the suppression sanction, the exclusionary rule, will force prosecutors and police to obey them.

There are many who would object to these requirements as unnecessary impediments in the path of aggressive law enforcement. I disagree. These requirements are the way we attempt to reach a balance between the claims of public safety and the claims of personal freedom in a nation that believes that government, including police, should not have limitless power. And only by demonstrating that our courts will not permit their violation can we underscore their importance or enforce them at all.

My views about the deterrence inherent in the close police-prosecution contacts with which I am most familiar are subject to the criticism that they do not describe the rule's impact on the uniformed officer on patrol or the agent on the street who is called upon to make split-second arrest and search decisions and who measures his success by arrests made, not convictions won.

There is force to this criticism. I don't doubt that the rule works less well in street encounters. But here, too, the rule is at work because of the enormous increase in police training and education about constitutional rights directly attributable to the exclusion sanction.

Federal agencies, subjected to the exclusionary rule since 1914, have traditionally devoted substantial time to such training. . . .

In my state, *Mapp* has been responsible for a virtual explosion in the amount and quantity of police training in the last twenty years. . . .

AN ASSESSMENT OF TWO PROPOSED MODIFICATIONS

The Good Faith Exception

Critics of the rule frequently complain that courts exclude illegally seized evidence even though the officers who seized it believed in good faith that their conduct was lawful. . . . The Task Force proposes that "in general, evidence should not be excluded from a criminal proceeding if it has been obtained by an officer acting in the reasonable, good faith belief that it was in conformity to the Fourth Amendment to the Constitution." Proposed legislation reflects this point of view, although it

seems to admit illegally seized evidence that even the "reasonable, good faith" rule would bar.

As several commentators have pointed out, the "good faith" test puts a premium on "police ignorance."[6] If the "good faith" loophole is established, pretrial hearings in every Fourth Amendment case will test the "reasonableness" of the police officer's ignorance of the Fourth Amendment's commands. Often (*very* often, given the antipathy of some judges toward the rule and public pressure to jettison it) the officer's ignorance will be found to have been "reasonable." And each time unconstitutional behavior is thus excused, the benchmarks of Fourth Amendment compliance will drop another notch. It is worth asking what impact such an exemption will have on the courses and classroom hours police will devote to the constitutional limitations on their profession if, in fact, it pays to be ignorant. Introduction of the "reasonable good faith" test would, I fear, depress Fourth Amendment compliance to the level of our tolerance for the lowest standards of the least informed officer. And there would be no incentive to do better. . . .

Advocates of the good faith [exception] . . . sometimes forget that the Fourth Amendment proscribes only "unreasonable" searches and seizures. Probable cause is supposed to turn on "the factual and practical considerations of everyday life on which reasonable and prudent men, not legal technicians, act." Scores of Supreme Court pronouncements make it clear that practical, common sense standards, including appreciation of the alternatives available to officers at the scene of crimes, are the guideposts for interpreting the Fourth Amendment's standard of reasonableness.

When courts fail to honor these standards of reasonableness, application of the exclusionary rule works to keep out evidence that should have been admitted. But once a violation of the constitutional commands of the Fourth Amendment is found, it would do serious violence to the deterrent purposes and societal values of the rule to say that, on the one hand, the police conduct is unconstitutional but that, on the other hand, the evidence it produced should be admitted anyway. The cure for such wrong interpretations of the Fourth Amendment is a better, sounder, and more practical judiciary. One does not have to agree with every court that ever found a search unconstitutional under the Fourth Amendment in order to believe, as I do, that the exclusionary rule is a sound remedy.

A Tort Remedy as Substitute for the Exclusionary Rule

[The proposal] which abolishes the exclusionary rule and replaces it with an exclusive tort remedy against the United States for damages resulting from a search conducted in violation of the Fourth Amendment is an open invitation to official lawlessness. If it were law, a federal officer could engage in essentially risk-free violations of the Constitution. If he engaged in an unconstitutional search, the risk of personal liability would not increase no matter how offensive the search might be. . . .

[Such a law might] even prompt agencies to encourage constitutional shortcuts by their employees. An agency that now may be required to spend time and money to put together a lawful case, including probable cause to search for incriminating evidence, might well seek to economize by accepting the fixed risk of a . . . "fine" assuming suit is even brought. . . . The agency might well conclude that the cost of doing business would be cheap at twice the price. Thus important constitutional protections are demeaned and budget cutting is brought to unaccustomed lengths.

Supplementing the deterrent effect of the exclusionary rule in federal court by a tort remedy against the United States makes sense. It would open an avenue of compensation for innocent victims of government misconduct not afforded by the suppression remedy. But to abolish the exclusionary rule and to replace it with a system that would fund unconstitutional conduct is to stand the Fourth Amendment on its head.

THE IMPERATIVE OF JUDICIAL INTEGRITY

It is sometimes said that the exclusionary rule breeds disrespect for the law because it suppresses the truth and permits crime to go unpunished. I believe that abolition of the rule would be far more destructive of respect for law. When an American court admits evidence obtained in violation of the Constitution, it is not merely permitting the truth to be heard. It is

inescapably condoning, validating, even welcoming, the illegality that produced it. It becomes part of that illegality. It paints a portrait of hypocrisy in a nation that professes to believe in the rule of law and whose courts, in the words of Madison, are to be the great "bulwarks" and "guardians" of our liberties. And the admission of such evidence is dangerous. As Justice Brandeis put it:

> In a government of laws, existence of the government will be imperilled if it fails to observe the law scrupulously. Our Government is the potent, the omnipresent teacher. For good or for ill, it reaches the whole people by its example. Crime is contagious. If the Government becomes a lawbreaker, it breeds contempt for law; it invites every man to become a law unto himself; it invites anarchy. To declare that in the administration of the criminal law the end justifies the means—to declare that the government may commit crimes in order to secure the conviction of a private criminal—would bring terrible retribution.

Law enforcement, in particular, needs public trust and respect for its authority in order to do its job effectively and safely. Official lawlessness destroys that trust, poisons police relations with the citizenry, and thus adds immeasurably to law enforcement's burdens.

It is easy to salute the liberties of the Bill of Rights in the abstract. But these freedoms have a price. It is difficult to remember, but we must never forget, that we cannot apply them selectively. Only insofar as we permit their effective exercise by the guilty will they remain strong protection for the innocent. Rights atrophy with disuse. They must be used not only in times of calm but in times of passion and fear as well.

In the play *A Man for All Seasons,* an account of the martyrdom of Saint Thomas More, More warns a zealot against "cut[ting] a great road through the law to get after the Devil."

"When the last law was down, and the Devil turned round on you, where would you hide," he asks, "the laws all being flat?" "Yes," More adds, "I'd give the Devil benefit of law, for my own safety's sake."

In this time of great passion about crime we should be extremely careful, for our own safety's sake, not to let our zeal to "get after the Devil" lead us to cut a great road through the Bill of Rights.

NOTES

1. *Mapp v. Ohio,* 367 U.S. 643 (1961).
2. Malcolm Wilkey, *Enforcing the Fourth Amendment by Alternatives to the Exclusionary Rule* (Washington, D.C.: National Legal Center for the Public Interest, March 1982).
3. Wilkey, *Enforcing the Fourth Amendment.*
4. Wayne LaFave, *Search and Seizure:* 1981 *Supplement* (St. Paul, Minn.: West Publishing Co., 1981), 1.2 n.9.
5. Yale Kamisar, "Defense of the Exclusionary Rule," *Crim. L. Bull.* 15 (1979): 13.
6. Ibid., p. 35, n.

REVIEW AND DISCUSSION QUESTIONS

1. Wilkey enumerates twelve costs of the exclusionary rule. Assess each of these. Which are the most important? Do any of them seem exaggerated?
2. What do you see as the strongest argument against the exclusionary rule? As the strongest argument for the exclusionary rule?
3. What three alternatives to the exclusionary rule does Wilkey propose? Are these likely to be as effective? How persuasive is Sachs's critique of the "good faith" exception and tort remedies?
4. Wilkey and Sachs disagree about the effectiveness of the exclusionary rule as a deterrent to police illegality and about whether some alternative method might protect our Fourth Amendment rights equally well but at a lower cost. Whose position do you find the most convincing and why?

Interrogation and the Right to Counsel

Miranda v. Arizona; Brewer v. Williams; and Rhode Island v. Innis

These three cases are among the Supreme Court's most important deliberations on criminal procedure and the rights of defendants. In *Miranda,* the Court held that procedural safeguards must be employed to respect the constitutional privilege against self-incrimination and, in particular, to inform accused persons of their right to remain silent—the now famous "Miranda warning." In *Brewer,* the Court debated the related issue of a defendant's right to counsel. Although three justices dissented from the decision, the majority determined that Robert Williams's conviction for murder was based on evidence obtained in violation of his constitutional right to the assistance of counsel. In the *Innis* case, the Court undertook to define and delimit the meaning of "interrogation." Although the Supreme Court of Rhode Island had relied on the U.S. Supreme Court's *Miranda* and *Brewer* decisions to overturn Innis's conviction, the U.S. Supreme Court, despite a vigorous dissent by Justice Stevens, decided that Innis had not been illegally interrogated.

Miranda v. Arizona

[*Editors' Note: The facts of this case were summarized in Justice Harlan's dissent from the majority opinion, as follows:*]

[I]t may make the analysis more graphic to consider the actual facts of one of the four cases reversed by the Court. *Miranda v. Arizona* serves best, being neither the hardest nor easiest of the four under the Court's standards.

On March 3, 1963, an 18-year-old girl was kidnapped and forcibly raped near Phoenix, Arizona. Ten days later, on the morning of March 13, petitioner Miranda was arrested and taken to the police station. At this time Miranda was 23 years old, indigent, and educated to the extent of completing half the ninth grade. He had "an emotional illness" of the schizophrenic type, according to the doctor who eventually examined him; the doctor's report also stated that Miranda was "alert and oriented as to time, place, and person," intelligent within normal limits, competent to stand trial, and sane within the legal definition. At the police station, the victim picked Miranda out of a line-up, and two officers then took him into a separate room to interrogate him, starting about 11:30 a.m. Though at first denying his guilt, within a short time Miranda gave a detailed oral confession and then wrote out in his own hand and signed a brief statement admitting and describing the crime. All this was accomplished in two hours or less without any force, threats or promises and—I will assume this though the record is uncertain . . . without any effective warnings at all.

Mr. Chief Justice Warren delivered the opinion of the court: The cases before us raise questions which go to the roots of our concepts of American criminal jurisprudence: the restraints society must observe consistent with the Federal Constitution in prosecuting individuals for crime. More specifically, we deal with the admissibility of statements obtained from an individual who is subjected to custodial police interrogation and the necessity for procedures which assure that the individual is accorded his privilege under the Fifth Amendment to the Constitution not to be compelled to incriminate himself. . . .

Our holding will be spelled out with some specificity in the pages which follow but briefly

384 U.S. 436 (1966); 430 U.S. 378 (1976); and 446 U.S. 291 (1979), respectively. Some footnotes and citations omitted.

stated it is this: the prosecution may not use statements, whether exculpatory or inculpatory, stemming from custodial interrogation of the defendant unless it demonstrates the use of procedural safeguards effective to secure the privilege against self-incrimination. By custodial interrogation, we mean questioning initiated by law enforcement officers after a person has been taken into custody or otherwise deprived of his freedom of action in any significant way. As for the procedural safeguards to be employed, unless other fully effective means are devised to inform accused persons of their right of silence and to assure a continuous opportunity to exercise it, the following measures are required. Prior to any questioning, the person must be warned that he has a right to remain silent, that any statement he does make may be used as evidence against him, and that he has a right to the presence of an attorney, either retained or appointed. The defendant may waive effectuation of these rights, provided the waiver is made voluntarily, knowingly and intelligently. If, however, he indicates in any manner and at any stage of the process that he wishes to consult with an attorney before speaking there can be no questioning. Likewise, if the individual is alone and indicates in any manner that he does not wish to be interrogated, the police may not question him. The mere fact that he may have answered some questions or volunteered some statements on his own does not deprive him of the right to refrain from answering any further inquiries until he has consulted with an attorney and thereafter consents to be questioned.

1

The constitutional issue we decide in each of these cases is the admissibility of statements obtained from a defendant questioned while in custody or otherwise deprived of his freedom of action in any significant way. In each, the defendant was questioned by police officers, detectives, or a prosecuting attorney in a room in which he was cut off from the outside world. In none of these cases was the defendant given a full and effective warning of his rights at the outset of the interrogation process. In all the cases, the questioning elicited oral admissions, and in three of them, signed statements as well which were admitted at their trials. They all thus share salient features—incommunicado interrogation of individuals in a police-dominated atmosphere, resulting in self-incriminating statements without full warnings of constitutional rights.

An understanding of the nature and setting of this in-custody interrogation is essential to our decisions today. The difficulty in depicting what transpires at such interrogations stems from the fact that in this country they have largely taken place incommunicado. From extensive factual studies undertaken in the early 1930s, including the famous Wickersham Report to Congress by a Presidential Commission, it is clear that police violence and the "third degree" flourished at that time. In a series of cases decided by this Court long after these studies, the police resorted to physical brutality—beating, hanging, whipping—and to sustained and protracted questioning incommunicado in order to extort confessions. The Commission on Civil Rights in 1961 found much evidence to indicate that "some policemen still resort to physical force to obtain confessions," 1961 Comm'n on Civil Rights Rep., Justice, pt. 5, 17. The use of physical brutality and violence is not, unfortunately, relegated to the past or to any part of the country. Only recently in Kings County, New York, the police brutally beat, kicked and placed lighted cigarette butts on the back of a potential witness under interrogation for the purpose of securing a statement incriminating a third party.[1]

The examples given above are undoubtedly the exception now, but they are sufficiently widespread to be the object of concern. Unless a proper limitation upon custodial interrogation is achieved—such as these decisions will advance—there can be no assurance that practices of this nature will be eradicated in the foreseeable future. The conclusion of the Wickersham Commission Report, made over 30 years ago, is still pertinent:

> To the contention that the third degree is necessary to get the facts, the reporters aptly reply in the language of the present Lord Chancellor of England (Lord Sankey): "It is not admissible to do a great right by doing a little wrong. . . . It is not sufficient to do justice by obtaining a proper result by irregular or

> improper means." Not only does the use of the third degree involve a flagrant violation of law by the officers of the law, but it involves also the dangers of false confessions, and it tends to make police and prosecutors less zealous in the search for objective evidence. As the New York prosecutor quoted in the report said, "It is a short cut and makes the police lazy and unenterprising." Or, as another official quoted remarked: "If you use your fists, you are not so likely to use your wits." We agree with the conclusion expressed in the report, that "The third degree brutalizes the police, hardens the prisoner against society, and lowers the esteem in which the administration of justice is held by the public." IV National Commission on Law Observance and Enforcement, Report on Lawlessness in Law Enforcement 5 (1931).

Again we stress that the modern practice of in-custody interrogation is psychologically rather than physically oriented. . . . Interrogation still takes place in privacy. Privacy results in secrecy and this in turn results in a gap in our knowledge as to what in fact goes on in the interrogation rooms. A valuable source of information about present police practices, however, may be found in various police manuals and texts which document procedures employed with success in the past, and which recommend various other effective tactics.[2] . . .

The officers are told by the manuals that the "principal psychological factor contributing to a successful interrogation is *privacy*—being alone with the person under interrogation."[3] The efficacy of this tactic has been explained as follows:

> If at all practicable, the interrogation should take place in the investigator's office or at least in a room of his own choice. The subject should be deprived of every psychological advantage. In his own home he may be confident, indignant, or recalcitrant. He is more keenly aware of his rights and more reluctant to tell of his indiscretions or criminal behavior within the walls of his home. Moreover his family and other friends are nearby, their presence lending moral support. In his own office, the investigator possesses all the advantages. The atmosphere suggests the invincibility of the forces of the law.[4]

To highlight the isolation and unfamiliar surroundings, the manuals instruct the police to display an air of confidence in the suspect's guilt and from outward appearance to maintain only an interest in confirming certain details. The guilt of the subject is to be posited as a fact. The interrogator should direct his comments toward the reasons why the subject committed the act, rather than court failure by asking the subject whether he did it. Like other men, perhaps the subject has had a bad family life, had an unhappy childhood, had too much to drink, had an unrequited desire for women. The officers are instructed to minimize the moral seriousness of the offense, to cast blame on the victim or on society. These tactics are designed to put the subject in a psychological state where his story is but an elaboration of what the police purport to know already—that he is guilty. Explanations to the contrary are dismissed and discouraged.

The texts thus stress that the major qualities an interrogator should possess are patience and perseverance. One writer describes the efficacy of these characteristics in this manner:

> In the preceding paragraphs emphasis has been placed on kindness and stratagems. The investigator will, however, encounter many situations where the sheer weight of his personality will be the deciding factor. Where emotional appeals and tricks are employed to no avail, he must rely on an oppressive atmosphere of dogged persistence. He must interrogate steadily and without relent, leaving the subject no prospect of surcease. He must dominate his subject and overwhelm him with his inexorable will to obtain the truth. He should interrogate for a spell of several hours pausing only for the subject's necessities in acknowledgment of the need to avoid a charge of duress that can be technically substantiated. In a serious case, the interrogation may continue for days, with the required intervals for food and sleep, but with no respite from the atmosphere of domination. It is possible in this way to induce the subject to talk without resorting to duress or coercion. The method should be used only when the guilt of the subject appears highly probable.[5]

. . . The interrogators sometimes are instructed to induce a confession out of trickery. The technique here is quite effective in crimes which require identification or which run in series. In the identification situation, the interrogator may take a break in his questioning to place the subject among

a group of men in a line-up. "The witness or complainant (previously coached, if necessary) studies the line-up and confidently points out the subject as the guilty party."[6] Then the questioning resumes "as though there were now no doubt about the guilt of the subject." A variation on this technique is called the "reverse line-up":

> The accused is placed in a line-up, but this time he is identified by several fictitious witnesses or victims who associated him with different offenses. It is expected that the subject will become desperate and confess to the offense under investigation in order to escape from the false accusations.[7]

. . . In the event that the subject wishes to speak to a relative or an attorney, the following advice is tendered:

> [T]he interrogator should respond by suggesting that the subject first tell the truth to the interrogator himself rather than get anyone else involved in the matter. If the request is for an attorney, the interrogator may suggest that the subject save himself or his family the expense of any such professional service, particularly if he is innocent of the offense under investigation. The interrogator may also add, "Joe, I'm only looking for the truth, and if you're telling the truth, that's it. You can handle this by yourself."[8]

From these representative samples of interrogation techniques, the setting prescribed by the manuals and observed in practice becomes clear. In essence, it is this: To be alone with the subject is essential to prevent distraction and to deprive him of any outside support. The aura of confidence in his guilt undermines his will to resist. He merely confirms the preconceived story the police seek to have him describe. Patience and persistence, at times relentless questioning, are employed. To obtain a confession, the interrogator must "patiently maneuver himself or his quarry into a position from which the desired objective may be attained."[9] When normal procedures fail to produce the needed result, the police may resort to deceptive stratagems such as giving false legal advice. It is important to keep the subject off balance, for example, by trading on his insecurity about himself or his surroundings. The police then persuade, trick, or cajole him out of exercising his constitutional rights.

Even without employing brutality, the "third degree" or the specific stratagems described above, the very fact of custodial interrogation exacts a heavy toll on individual liberty and trades on the weakness of individuals. . . .

2

. . . Those who framed our Constitution and the Bill of Rights were ever aware of subtle encroachments on individual liberty. They knew that "illegitimate and unconstitutional practices get their first footing . . . by silent approaches and slight deviations from legal modes of procedure." *Boyd v. United States,* 116 U.S. 616, 635 (1886).

. . . We may view the historical development of the privilege [against self-incrimination] as one which groped for the proper scope of governmental power over the citizen. As a "noble principle often transcends its origins," the privilege has come rightfully to be recognized in part as an individual's substantive right, a "right to a private enclave where he may lead a private life. That right is the hallmark of our democracy." *United States v. Grunewald,* 233 F.2d. . . . [The] constitutional foundation underlying the privilege is the respect a government—state or federal—must accord to the dignity and integrity of its citizens. To maintain a "fair state-individual balance," to require the government "to shoulder the entire load," . . . to respect the inviolability of the human personality, our accusatory system of criminal justice demands that the government seeking to punish an individual produce the evidence against him by its own independent labors, rather than by the cruel, simple expedient of compelling it from his own mouth. . . . In sum, the privilege is fulfilled only when the person is guaranteed the right "to remain silent unless he chooses to speak in the unfettered exercise of his own will." *Malloy v. Hogan,* 378 U.S. 1, 8 (1964).

. . . We are satisfied that all the principles embodied in the privilege apply to informal compulsion exerted by law-enforcement officers during in-custody questioning. An individual swept from familiar surroundings into police custody, surrounded by antagonistic forces, and subjected to the techniques of persuasion described above cannot be otherwise than under compulsion to speak.

As a practical matter, the compulsion to speak in the isolated setting of the police station may well be greater than in courts or other official investigations, where there are often impartial observes to guard against intimidation or trickery. . . .

3

Today, then, there can be no doubt that the Fifth Amendment privilege is available outside of criminal court proceedings and serves to protect persons in all settings in which their freedom of action is curtailed in any significant way from being compelled to incriminate themselves. We have concluded that without proper safeguards the process of in-custody interrogation of persons suspected or accused of crime contains inherently compelling pressures which work to undermine the individual's will to resist and to compel him to speak where he would not otherwise do so freely. In order to combat these pressures and to permit a full opportunity to exercise the privilege against self-incrimination, the accused must be adequately and effectively apprised of his rights and the exercise of those rights must be fully honored. . . .

The principles announced today deal with the protection which must be given to the privilege against self-incrimination when the individual is first subjected to police interrogation while in custody at the station or otherwise deprived of his freedom of action in any significant way. It is at this point that our adversary system of criminal proceedings commences, distinguishing itself at the outset from the inquisitorial system recognized in some countries. Under the system of warnings we delineate today or under any other system which may be devised and found effective, the safeguards to be erected about the privilege must come into play at this point.

Our decision is not intended to hamper the traditional function of police officers in investigating crime. . . . When an individual is in custody on probable cause, the police may, of course, seek out evidence in the field to be used at trial against him. Such investigation may include inquiry of persons not under restraint. General on-the-scene questioning as to facts surrounding a crime or other general questioning of citizens in the fact-finding process is not affected by our holding. It is an act of responsible citizenship for individuals to give whatever information they may have to aid in law enforcement. In such situations the compelling atmosphere inherent in the process of in-custody interrogation is not necessarily present.

In dealing with statements obtained through interrogation, we do not purport to find all confessions inadmissible. Confessions remain a proper element in law enforcement. Any statement given freely and voluntarily without any compelling influences is, of course, admissible in evidence. The fundamental import of the privilege while an individual is in custody is not whether he is allowed to talk to the police without the benefit of warnings and counsel, but whether he can be interrogated. There is no requirement that police stop a person who enters a police station and states that he wishes to confess to a crime, or a person who calls the police to offer a confession or any other statement he desires to make. Volunteered statements of any kind are not barred by the Fifth Amendment and their admissibility is not affected by our holding today. . . .

To summarize, we hold that when an individual is taken into custody or otherwise deprived of his freedom by the authorities in any significant way and is subjected to questioning, the privilege against self-incrimination is jeopardized. Procedural safeguards must be employed to protect the privilege, and unless other fully effective means are adopted to notify the person of his right of silence and to assure that the exercise of the right will be scrupulously honored, the following measures are required. He must be warned prior to any questioning that he has the right to remain silent, that anything he says can be used against him in a court of law, that he has the right to the presence of an attorney, and that if he cannot afford an attorney one will be appointed for him prior to any questioning if he so desires. Opportunity to exercise these rights must be afforded to him throughout the interrogation. After such warnings have been given, and such opportunity afforded him, the individual may knowingly and intelligently waive these rights and agree to answer questions or make a statement. But unless and until such warnings and waiver are demonstrated by the prosecution at trial, no evidence obtained as a result of interrogation can be used against him.

NOTES

1. In addition, see *People v. Wakat* . . . (defendant suffering from broken bones, multiple bruises and injuries sufficiently serious to require eight months' medical treatment after being manhandled by five policemen); *Kier v. State* . . . (police doctor told accused, who was strapped to a chair completely nude, that he proposed to take hair and skin scrapings from anything that looked like blood or sperm from various parts of his body); *Bruner v. People* . . . (defendant held in custody over two months, deprived of food for 15 hours, forced to submit to a lie detector test when he wanted to go to the toilet); *People v. Matlock* . . . (defendant questioned incessantly over an evening's time, made to lie on cold board and to answer questions whenever it appeared he was getting sleepy).
2. The manuals quoted in the text following are the most recent and representative of the texts currently available.
3. Inbau and Reid, *Criminal Interrogation and Confessions* (1962), at 1.
4. O'Hara, *Fundamentals of Criminal Investigation* (1956), at 99.
5. O'Hara, *supra,* at 112.
6. O'Hara, *supra,* at 105–106.
7. *Id.,* at 106.
8. Inbau and Reid, *supra,* at 112.
9. Inbau and Reid, *Lie Detection and Criminal Interrogation* 185 (3d ed. 1953).

Brewer v. Williams

Mr. Justice Stewart delivered the opinion of the court: An Iowa trial jury found the respondent, Robert Williams, guilty of murder. The judgment of conviction was affirmed in the Iowa Supreme Court by a closely divided vote. In a subsequent habeas corpus proceeding a Federal District Court ruled that under the United States Constitution Williams is entitled to a new trial, and a divided Court of Appeals for the Eighth Circuit agreed. The question before us is whether the District Court and the Court of Appeals were wrong.

1

On the afternoon of December 24, 1968, a 10-year-old girl named Pamela Powers went with her family to the YMCA in Des Moines, Iowa, to watch a wrestling tournament in which her brother was participating. When she failed to return from a trip to the washroom, a search for her began. The search was unsuccessful.

Robert Williams, who had recently escaped from a mental hospital, was a resident of the YMCA. Soon after the girl's disappearance Williams was seen in the YMCA lobby carrying some clothing and a large bundle wrapped in a blanket. He obtained help from a 14-year-old boy in opening the street door of the YMCA and the door to his automobile parked outside. When Williams placed the bundle in the front seat of his car the boy "saw two legs in it and they were skinny and white." Before anyone could see what was in the bundle Williams drove away. His abandoned car was found the following day in Davenport, Iowa, roughly 160 miles east of Des Moines. A warrant was then issued in Des Moines for his arrest on a charge of abduction.

On the morning of December 26, a Des Moines lawyer named Henry McKnight went to the Des Moines police station and informed the officers present that he had just received a long-distance call from Williams, and that he had advised Williams to turn himself in to the Davenport police. Williams did surrender that morning to the police in Davenport, and they booked him on the charge specified in the arrest warrant and gave him the warnings required by *Miranda v. Arizona,* 384 U.S. 436. The Davenport police then telephoned their counter-parts in Des Moines to inform them that Williams had surrendered. McKnight, the lawyer, was still at the Des Moines police headquarters, and Williams conversed with McKnight on the telephone. In the presence of the Des Moines chief of police and a police detective named Leaming, McKnight advised Williams that Des Moines police officers would be driving to Davenport to pick him up, that the officers would not interrogate him or mistreat him, and that Williams was not to talk to the officers about Pamela Powers until after consulting with McKnight upon his return to Des Moines.

As a result of these conversations, it was agreed between McKnight and the Des Moines police officials that Detective Leaming and a fellow officer would drive to Davenport to pick up Williams, that they would bring him directly back to Des Moines, and that they would not question him during the trip.

In the meantime Williams was arraigned before a judge in Davenport on the outstanding arrest warrant. The judge advised him of his *Miranda* rights and committed him to jail. Before leaving the courtroom, Williams conferred with a lawyer named Kelly, who advised him not to make any statements until consulting with McKnight back in Des Moines.

Detective Leaming and his fellow officer arrived in Davenport about noon to pick up Williams and return him to Des Moines. Soon after their arrival they met with Williams and Kelly, who, they understood, was acting as Williams's lawyer. Detective Leaming repeated the *Miranda* warnings, and told Williams:

> [W]e both know that you're being represented here by Mr. Kelly and you're being represented by Mr. McKnight in Des Moines, and . . . I want you to remember this because we'll be visiting between here and Des Moines.

Williams then conferred again with Kelly alone, and after this conference Kelly reiterated to Detective Leaming that Williams was not to be questioned about the disappearance of Pamela Powers until after he had consulted with McKnight back in Des Moines. When Leaming expressed some reservations, Kelly firmly stated that the agreement with McKnight was to be carried out—that there was to be no interrogation of Williams during the automobile journey to Des Moines. Kelly was denied permission to ride in the police car back to Des Moines with Williams and the two officers.

The two detectives, with Williams in their charge, then set out on the 160-mile drive. At no time during the trip did Williams express a willingness to be interrogated in the absence of an attorney. Instead, he stated several times that "[w]hen I get to Des Moines and see Mr. McKnight, I am going to tell you the whole story." Detective Leaming knew that Williams was a former mental patient, and knew also that he was deeply religious.

The detective and his prisoner soon embarked on a wide-ranging conversation covering a variety of topics, including the subject of religion. Then, not long after leaving Davenport and reaching the interstate highway, Detective Leaming delivered what has been referred to in the briefs and oral arguments as the "Christian burial speech." Addressing Williams as "Reverend," the detective said:

> I want to give you something to think about while we're traveling down the road. . . . Number one, I want you to observe the weather conditions, it's raining, it's sleeting, it's freezing, driving is very treacherous, visibility is poor, it's going to be dark early this evening. They are predicting several inches of snow for tonight, and I feel that you yourself are the only person that knows where this little girl's body is, that you yourself have only been there once, and if you get a snow on top of it you yourself may be unable to find it. And, since we will be going right past the area on the way into Des Moines, I feel that we could stop and locate the body, that the parents of this little girl should be entitled to a Christian burial for the little girl who was snatched away from them on Christmas [E]ve and murdered. And I feel we should stop and locate it on the way in rather than waiting until morning and trying to come back out after a snow storm and possibly not being able to find it at all.

Williams asked Detective Leaming why he thought their route to Des Moines would be taking them past the girl's body, and Leaming responded that he knew the body was in the area of Mitchellville—a town they would be passing on the way to Des Moines.[1] Leaming then stated: "I do not want you to answer me. I don't want to discuss it any further. Just think about it as we're riding down the road."

. . . The car continued towards Des Moines and as it approached Mitchellville, Williams said that he would show the officers where the body was. He then directed the police to the body of Pamela Powers.

Williams was indicted for first-degree murder. . . .

The evidence in question was introduced over counsel's continuing objection at the subsequent trial. The jury found Williams guilty of murder. . . .

2

. . . [T]here is no need to review in this case the doctrine of *Miranda v. Arizona,* a doctrine designed to secure the constitutional privilege against compulsory self-incrimination. . . . For it is clear that the judgment before us must in any event be affirmed upon the ground that Williams was deprived of a different constitutional right—the right to the assistance of counsel.

This right, guaranteed by the Sixth and Fourteenth Amendments, is indispensable to the fair administration of our adversary system of criminal justice.

. . . Whatever else it may mean, the right to counsel granted by the Sixth and Fourteenth Amendments means at least that a person is entitled to the help of a lawyer at or after the time that judicial proceedings have been initiated against him—"whether by way of formal charge, preliminary hearing, indictment, information, or arraignment." . . .

There can be no doubt in the present case that judicial proceedings had been initiated against Williams before the start of the automobile ride from Davenport to Des Moines. A warrant had been issued for his arrest, he had been arraigned on that warrant before a judge in a Davenport courtroom, and he had been committed by the court to confinement in jail. The State does not contend otherwise.

There can be no serious doubt, either, that Detective Leaming deliberately and designedly set out to elicit information from Williams just as surely as—and perhaps more effectively than—if he had formally interrogated him. Detective Leaming was fully aware before departing for Des Moines that Williams was being represented in Davenport by Kelly and in Des Moines by McKnight. Yet he purposely sought during Williams' isolation from his lawyers to obtain as much incriminating information as possible. Indeed, Detective Leaming conceded as much when he testified at Williams' trial:

Q: In fact, Captain, whether he was a mental patient or not, you were trying to get all the information you could before be got to his lawyer; weren't you?

A: I was sure hoping to find out where that little girl was, yes, sir. . . .

Q: Well, I'll put it this way: You was [sic] hoping to get all the information you could before Williams got back to McKnight, weren't you?

A: Yes, sir.

The state courts clearly proceeded upon the hypothesis that Detective Leaming's "Christian burial speech" had been tantamount to interrogation. Both courts recognized that Williams had been entitled to the assistance of counsel at the time he made the incriminating statements. Yet no such constitutional protection would have come into play if there had been no interrogation. . . .

3

The Iowa courts . . . held . . . that he had waived that right during the course of the automobile trip from Davenport to Des Moines. The state trial court explained its determination of waiver as follows:

> The time element involved on the trip, the general circumstances of it, and more importantly the absence on the Defendant's part of any assertion of his right or desire not to give information absent the presence of his attorney, are the main foundations for the Court's conclusion that he voluntarily waived such right.

. . . After carefully reviewing the evidence, the District Court concluded:

> [U]nder the proper standards for determining waiver, there simply is no evidence to support a waiver. . . . [T]here is no affirmative indication . . . that [Williams] did waive his rights. . . . [T]he state courts' emphasis on the absence of a demand for counsel was not only legally inappropriate, but factually unsupportable as well, since Detective Leaming himself testified that [Williams], on several occasions during the trip, indicated that he would talk *after* he saw Mr. McKnight. . . . Moreover, the statements were obtained only after Detective Leaming's use of psychology on a person whom he knew to be deeply religious and an escapee from a mental hospital—with the specific intent to elicit incriminating statements. In the face of this evidence, the State has produced no

> affirmative evidence whatsoever to support its claim of waiver, and, a fortiori, it cannot be said that the State has met its "heavy burden" of showing a knowing and intelligent waiver of . . . Sixth Amendment rights. *Id.*, at 182–183.

The Court of Appeals approved the reasoning of the District Court. . . .

The District Court and the Court of Appeals were . . . correct in their understanding of the proper standard to be applied in determining the question of waiver as a matter of federal constitutional law—that it was incumbent upon the State to prove "an intentional relinquishment or abandonment of a known right or privilege." *Johnson v. Zerbst,* 304 U.S:, at 464. That standard has been reiterated in many cases. We have said that the right to counsel does not depend upon a request by the defendant, *Miranda v. Arizona,* 384 U.S., at 471, and that courts indulge in every reasonable presumption against waiver. . . .

[Williams's] statements while in the car that he would tell the whole story *after* seeing McKnight in Des Moines were the clearest expressions by Williams himself that he desired the presence of an attorney before an interrogation took place. But even before making these statements, Williams had effectively asserted his right to counsel by having secured attorneys at both ends of the automobile trip, both of whom, acting as his agents, had made clear to the police that no interrogation was to occur during the journey. Williams knew of that agreement and, particularly in view of his consistent reliance on counsel, there is no basis for concluding that he disavowed it.

Despite Williams's express and implicit assertions of his right to counsel, Detective Leaming proceeded to elicit incriminating statements from Williams. Leaming did not preface this effort by telling Williams that he had a right to the presence of a lawyer, and made no effort at all to ascertain whether Williams wished to relinquish that right. The circumstances of record in this case thus provide no reasonable basis for finding that Williams waived his right to the assistance of counsel.

The Court of Appeals did not hold, nor do we, that under the circumstances of this case Williams *could not,* without notice to counsel, have waived his rights under the Sixth and Fourteenth Amendments. It only held, as do we, that he did not.

4

The crime of which Williams was convicted was senseless and brutal, calling for swift and energetic action by the police to apprehend the perpetrator and gather evidence with which he could be convicted. No mission of law enforcement officials is more important. Yet "[d]isinterested zeal for the public good does not assure either wisdom or right in the methods it pursues." *Haley v. Ohio* . . . Although we do not lightly affirm the issuance of a writ of habeas corpus in this case, so clear a violation of the Sixth and Fourteenth Amendments as here occurred cannot be condoned. The pressures on state executive and judicial officers charged with the administration of the criminal law are great, especially when the crime is murder and the victim a small child. But it is precisely the predictability of those pressures that makes imperative a resolute loyalty to the guarantees that the Constitution extends to us all.

The judgment of the Court of Appeals is affirmed.

It is so ordered.

Mr. Chief Justice Burger, dissenting: The result in this case ought to be intolerable in any society which purports to call itself an organized society. It continues the Court—by the narrowest margin—on the much-criticized course of punishing the public for the mistakes and misdeeds of law enforcement officers, instead of punishing the officer directly, if in fact he is guilty of wrongdoing. It mechanically and blindly keeps reliable evidence from juries whether the claimed constitutional violation involves gross police misconduct or honest human error.

Williams is guilty of the savage murder of a small child; no member of the Court contends he is not. While in custody, and after no fewer than *five* warnings of his rights to silence and to counsel, he led police to the concealed body of his victim. The Court concedes Williams was not threatened or coerced and that he spoke and acted voluntarily

and with full awareness of his constitutional rights. In the face of all this, the Court now holds that because Williams was prompted by the detective's statement—not interrogation but a statement—the jury must not be told how the police found the body.

Today's holding fulfills Judge (later Mr. Justice) Cardozo's grim prophecy that someday some court might carry the exclusionary rule to the absurd extent that its operative effect would exclude evidence relating to the body of a murder victim because of the means by which it was found. In so ruling the Court regresses to playing a grisly game of "hide and seek," once more exalting the sporting theory of criminal justice which has been experiencing a decline in our jurisprudence. With Justices White, Blackmun, and Rehnquist, I categorically reject the remarkable notion that the police in this case were guilty of unconstitutional misconduct, or any conduct justifying the bizarre result reached by the Court.

Mr. Justice White, with whom Mr. Justice Blackmun and Mr. Justice Rehnquist join, dissenting: ... The majority simply finds that no waiver was *proved* in this case. I disagree. That respondent knew of his right not to say anything to the officers without advice and presence of counsel is established on this record to a moral certainty. He was advised of the right by three officials of the State—telling at least one that he understood the right—and by two lawyers. Finally, he further demonstrated his knowledge of the right by informing the police that he would tell them the story in the presence of McKnight when they arrived in Des Moines. The issue in this case, then, is whether respondent relinquished that right intentionally.

Respondent relinquished his right not to talk to the police about his crime when the car approached the place where he had hidden the victim's clothes. Men usually intend to do what they do, and there is nothing in the record to support the proposition that respondent's decision to talk was anything but an exercise of his own free will. Apparently, without any prodding from the officers, respondent—who had earlier said that he would tell the whole story when he arrived in Des Moines—spontaneously changed his mind about the timing of his disclosures when the car approached the places where he had hidden the evidence. However, even if his statements were influenced by Detective Leaming's above-quoted statement, respondent's decision to talk in the absence of counsel can hardly be viewed as the product of an overborne will. The statement by Leaming was not coercive; it was accompanied by a request that respondent not respond to it; and it was delivered hours before respondent decided to make any statement. Respondent's waiver was thus knowing and intentional.

The majority's contrary conclusion seems to rest on the fact that respondent "asserted" his right to counsel by retaining and consulting with one lawyer and by consulting with another. How this supports the conclusion that respondent's later relinquishment of his right not to talk in the absence of counsel was unintentional is a mystery. The fact that respondent consulted with counsel on the question whether he should talk to the police in counsel's absence makes his later decision to talk in counsel's absence *better* informed and, if anything, more intelligent.

... Waiver is shown whenever the facts establish that an accused knew of a right and intended to relinquish it. Such waiver, even if not express, was plainly shown here. The only other conceivable basis for the majority's holding is the implicit suggestion ... that the right ... involved in *Miranda v. Arizona* ... is a right not to be *asked* any questions in counsel's absence rather than a right not to *answer* any questions in counsel's absence, and that the right not to be *asked* questions must be waived *before* the questions are asked. Such wafer-thin distinctions cannot determine whether a guilty murderer should go free. The only conceivable purpose for the presence of counsel during questioning is to protect an accused from making incriminating *answers.* Questions, unanswered, have no significance at all.

NOTE

1. The fact of the matter, of course, was that Detective Leaming possessed no such knowledge.

Rhode Island v. Innis

Mr. Justice Stewart delivered the opinion of the court: In *Miranda v. Arizona* the Court held that, once a defendant in custody asks to speak with a lawyer, all interrogation must cease until a lawyer is present. The issue in this case is whether the respondent was "interrogated" in violation of the standards promulgated in the *Miranda* opinion.

1

On the night of January 12, 1975, John Mulvaney, a Providence, R.I., taxicab driver, disappeared after being dispatched to pick up a customer. His body was discovered four days later buried in a shallow grave in Coventry, R.I. He had died from a shotgun blast aimed at the back of his head.

On January 17, 1975, shortly after midnight, the Providence police received a telephone call from Gerald Aubin, also a taxicab driver, who reported that he had just been robbed by a man wielding a sawed-off shotgun. . . . Aubin identified a picture of the same person. That person was the respondent. Shortly thereafter, the Providence police began a search of the Mount Pleasant area.

At approximately 4:30 a.m. on the same date, Patrolman Lovell, while cruising the streets of Mount Pleasant in a patrol car, spotted the respondent standing in the street facing him. When Patrolman Lovell stopped his car, the respondent walked towards it. Patrolman Lovell then arrested the respondent, who was unarmed, and advised him of his so-called *Miranda* rights. While the two men waited in the patrol car for other police officers to arrive, Patrolman Lovell did not converse with the respondent other than to respond to the latter's request for a cigarette.

Within minutes, Sergeant Sears arrived at the scene of the arrest, and he also gave the respondent the *Miranda* warnings. Immediately thereafter, Captain Leyden and other police officers arrived. Captain Leyden advised the respondent of his *Miranda* rights. The respondent stated that he understood those rights and wanted to speak with a lawyer. Captain Leyden then directed that the respondent be placed in a "caged wagon," a four-door police car with a wire screen mesh between the front and rear seats, and be driven to the central police station. Three officers, Patrolmen Gleckman, Williams, and McKenna, were assigned to accompany the respondent to the central station. . . .

While en route to the central station, Patrolman Gleckman initiated a conversation with Patrolman McKenna concerning the missing shotgun. As Patrolman Gleckman later testified:

A. At this point, I was talking back and forth with Patrolman McKenna stating that I frequent this area while on patrol and [that because a school for handicapped children is located nearby,] there's a lot of handicapped children running around in this area, and God forbid one of them might find a weapon with shells and they might hurt themselves. App. 43–44.

Patrolman McKenna apparently shared his fellow officer's concern:

A. I more or less concurred with him [Gleckman] that it was a safety factor and that we should, you know, continue to search for the weapon and try to find it. Id., at 53.

While Patrolman Williams said nothing, he overheard the conversation between the two officers:

A. He [Gleckman] said it would be too bad if the little—I believe he said a girl—would pick up the gun, maybe kill herself. Id., at 59.

The respondent then interrupted the conversation, stating that the officers should turn the car around so he could show them where the gun was located. At this point, Patrolman McKenna radioed back to Captain Leyden that they were returning to the scene of the arrest, and that the respondent would inform them of the location of the gun. At the time the respondent indicated that the officers should turn back, they had traveled no more than a mile, a trip encompassing only a few minutes.

The police vehicle then returned to the scene of the arrest where a search for the shotgun was in progress. There, Captain Leyden again advised the respondent of his *Miranda* rights. The respondent

replied that he understood those rights but that he "wanted to get the gun out of the way because of the kids in the area in the school." The respondent then led the police to a nearby field, where he pointed out the shotgun under some rocks by the side of the road.

On March 20, 1975, a grand jury returned an indictment charging the respondent with the kidnaping, robbery, and murder of John Mulvaney. Before trial, the respondent moved to suppress the shotgun and the statements he had made to the police regarding it. After an evidentiary hearing at which the respondent elected not to testify, the trial judge found that the respondent had been "repeatedly and completely advised of his *Miranda* rights." He further found that it was "entirely understandable that [the officers in the police vehicle] would voice their concern [for the safety of the handicapped children] to each other." The judge then concluded that the respondent's decision to inform the police of the location of the shotgun was "a waiver, clearly, and on the basis of the evidence that I have heard, and [*sic*] intelligent waiver, of his [*Miranda*] right to remain silent." Thus, without passing on whether the police officers had in fact "interrogated" the respondent, the trial court sustained the admissibility of the shotgun and testimony related to its discovery. That evidence was later introduced at the respondent's trial, and the jury returned a verdict of guilty on all counts.

On appeal, the Rhode Island Supreme Court, in a 3–2 decision, set aside the respondent's conviction. . . . Relying at least in part on this Court's decision in *Brewer v. Williams,* the court concluded that the respondent had invoked his *Miranda* right to counsel and that, contrary to *Miranda*'s mandate that, in the absence of counsel, all custodial interrogation then cease, the police officers in the vehicle had "interrogated" the respondent without a valid waiver of his right to counsel. It was the view of the state appellate court that, even though the police officers may have been genuinely concerned about the public safety and even though the respondent had not been addressed personally by the police officers, the respondent nonetheless had been subjected to "subtle coercion" that was the equivalent of "interrogation" within the meaning of the *Miranda* opinion. Moreover, contrary to the holding of the trial court, the appellate court concluded that the evidence was insufficient to support a finding of waiver. Having concluded that both the shotgun and testimony relating to its discovery were obtained in violation of the *Miranda* standards and therefore should not have been admitted into evidence, the Rhode Island Supreme Court held that the respondent was entitled to a new trial.

We granted certiorari to address for the first time the meaning of "interrogation" under *Miranda v. Arizona.*

2

. . . The starting point for defining "interrogation" in this context is, of course, the Court's *Miranda* opinion. There the Court observed that "[b]y custodial interrogation, we mean *questioning* initiated by law enforcement officers after a person has been taken into custody or otherwise deprived of his freedom of action in any significant way." This passage and other references throughout the opinion to "questioning" might suggest that the *Miranda* rules were to apply only to those police interrogation practices that involve express questioning of a defendant while in custody.

We do not, however, construe the *Miranda* opinion so narrowly. The concern of the Court in *Miranda* was that the "interrogation environment" created by the interplay of interrogation and custody would "subjugate the individual to the will of his examiner" and thereby undermine the privilege against compulsory self-incrimination. The police practices that evoked this concern included several that did not involve express questioning. For example, one of the practices discussed in *Miranda* was the use of line-ups in which a coached witness would pick the defendant as the perpetrator. This was designed to establish that the defendant was in fact guilty as a predicate for further interrogation. A variation on this theme discussed in *Miranda* was the so-called "reverse lineup" in which a defendant would be identified by coached witnesses as the perpetrator of a fictitious crime, with the object of inducing him to confess to the actual crime of which he was suspected in order to escape the false prosecution.

The Court in *Miranda* also included in its survey of interrogation practices the use of psychological ploys, such as to "posi[t]" "the guilt of the subject," to "minimize the moral seriousness of the offense," and "to cast blame on the victim or on society." It is clear that these techniques of persuasion, no less than express questioning, were thought, in a custodial setting, to amount to interrogation. . . .

We conclude that the *Miranda* safeguards come into play whenever a person in custody is subjected to either express questioning or its functional equivalent. That is to say, the term "interrogation" under *Miranda* refers not only to express questioning, but also to any words or actions on the part of the police (other than those normally attendant to arrest and custody) that the police should know are reasonably likely to elicit an incriminating response from the suspect. The latter portion of this definition focuses primarily upon the perceptions of the suspect, rather than the intent of the police. This focus reflects the fact that the *Miranda* safeguards were designed to vest a suspect in custody with an added measure of protection against coercive police practices, without regard to objective proof of the underlying intent of the police. A practice that the police should know is reasonably likely to evoke an incriminating response from a suspect thus amounts to interrogation. But, since the police surely cannot be held accountable for the unforeseeable results of their words or actions, the definition of interrogation can extend only to words or actions on the part of police officers that they *should have known* were reasonably likely to elicit an incriminating response.[1]

Turning to the facts of the present case, we conclude that the respondent was not "interrogated" with the meaning of *Miranda.* It is undisputed that the first prong of the definition of "interrogation" was not satisfied, for the conversation between Patrolmen Gleckman and McKenna included no express questioning of the respondent. Rather, that conversation was, at least in form, nothing more than a dialogue between the two officers to which no response from the respondent was invited.

Moreover, it cannot be fairly concluded that the respondent was subjected to the "functional equivalent" of questioning. It cannot be said, in short, that Patrolmen Gleckman and McKenna should have known that their conversation was reasonably likely to elicit an incriminating response from the respondent. There is nothing in the record to suggest that the officers were aware that the respondent was peculiarly susceptible to an appeal to his conscience concerning the safety of handicapped children. Nor is there anything in the record to suggest that the police knew that the respondent was unusually disoriented or upset at the time of his arrest.[2]

The case thus boils down to whether, in the context of a brief conversation, the officers should have known that the respondent would suddenly be moved to make a self-incriminating response. Given the fact that the entire conversation appears to have consisted of no more than a few offhand remarks, we cannot say that the officers should have known that it was reasonably likely that Innis would so respond. This is not a case where the police carried on a lengthy harangue in the presence of the suspect. Nor does the record support the respondent's contention that, under the circumstances, the officers' comments were particularly "evocative." It is our view, therefore, that the respondent was not subjected by the police to words or actions that the police should have known were reasonably likely to elicit an incriminating response from him.

The Rhode Island Supreme Court erred, in short, in equating "subtle compulsion" with interrogation. That the officers' comments struck a responsive chord is readily apparent. Thus, it may be said, as the Rhode Island Supreme Court did say, that the respondent was subjected to "subtle compulsion." But that is not the end of the inquiry. It must also be established that a suspect's incriminating response was the product of words or actions on the part of the police that they should have known were reasonably likely to elicit an incriminating response. This was not established in the present case.

Mr. Justice Stevens, dissenting:

1

. . . As the Court recognizes, *Miranda v. Arizona* makes it clear that, once respondent requested an attorney, he had an absolute right to have any type of interrogation cease until an attorney was

present. As it also recognizes, *Miranda* requires that the term "interrogation" be broadly construed to include "either express questioning or its functional equivalent." In my view any statement that would normally be understood by the average listener as calling for a response is the functional equivalent of a direct question, whether or not it is punctuated by a question mark. The Court, however, takes a much narrower view. It holds that police conduct is not the "functional equivalent" of direct questioning unless the police should have known that what they were saying or doing was likely to elicit an incriminating response from the suspect.[3] This holding represents a plain departure from the principles set forth in *Miranda.*

In *Miranda* the Court required the now-familiar warnings to be given to suspects prior to custodial interrogation in order to dispel the atmosphere of coercion that necessarily accompanies such interrogations. In order to perform that function effectively, the warnings must be viewed by both the police and the suspect as a correct and binding statement of their respective rights. Thus, if after being told that he has a right to have an attorney present during interrogation, a suspect chooses to cut off questioning until counsel can be obtained, his choice must be "scrupulously honored" by the police. . . . At the least this must mean that the police are prohibited from making deliberate attempts to elicit statements from the suspect. Yet the Court is unwilling to characterize all such attempts as "interrogation," noting only that "where a police practice is designed to elicit an incriminating response from the accused, it is unlikely that the practice will not also be one which the police should have known was reasonably likely to have that effect."

From the suspect's point of view, the effectiveness of the warnings depends on whether it appears that the police are scrupulously honoring his rights. Apparent attempts to elicit information from a suspect after he has invoked his right to cut off questioning necessarily demean that right and tend to reinstate the imbalance between police and suspect that the *Miranda* warnings are designed to correct. Thus, if the rationale for requiring those warnings in the first place is to be respected, any police conduct or statements that would appear to a reasonable person in the suspect's position to call for a response must be considered "interrogation."

In short, in order to give full protection to a suspect's right to be free from any interrogation at all, the definition of "interrogation" must include any police statement or conduct that has the same purpose or effect as a direct question. Statements that appear to call for a response from the suspect, as well as those that are designed to do so, should be considered interrogation. By prohibiting only those relatively few statements or actions that a police officer should know are likely to elicit an incriminating response, the Court today accords a suspect considerably less protection. Indeed, since I suppose most suspects are unlikely to incriminate themselves even when questioned directly, this new definition will almost certainly exclude every statement that is not punctuated with a question mark from the concept of "interrogation."

The difference between the approach required by a faithful adherence to *Miranda* and the stinted test applied by the Court today can be illustrated by comparing three different ways in which Officer Gleckman could have communicated his fears about the possible dangers posed by the shotgun to handicapped children. He could have:

> **1.** directly asked Innis: Will you please tell me where the shotgun is so we can protect handicapped schoolchildren from danger?
>
> **2.** announced to the other officers in the wagon: If the man sitting in the back seat with me should decide to tell us where the gun is, we can protect handicapped children from danger, or
>
> **3.** stated to the other officers: It would be too bad if a little handicapped girl would pick up the gun that this man left in the area and maybe kill herself.

In my opinion, all three of these statements should be considered interrogation because all three appear to be designed to elicit a response from anyone who in fact knew where the gun was located. Under the Court's test, on the other hand, the form of the statements would be critical. The third statement would not be interrogation because in the Court's view there was no reason for Officer Gleckman to believe that Innis was susceptible to this type of an implied appeal; therefore, the statement would not be reasonably likely to elicit an incriminating response. Assuming that this is true, . . . then it seems to me that the first two

statements, which would be just as unlikely to elicit such a response, should also not be considered interrogation. But, because the first statement is clearly an express question, it *would* be considered interrogation under the Court's test. The second statement, although just as clearly a deliberate appeal to Innis to reveal the location of the gun, would presumably not be interrogation because (a) it was not in form a direct question and (b) it does not fit within the "reasonably likely to elicit an incriminating response" category that applies to indirect interrogation.

As this example illustrates, the Court's test creates an incentive for police to ignore a suspect's invocation of his rights in order to make continued attempts to extract information from him. If a suspect does not appear to be susceptible to a particular type of psychological pressure, the police are apparently free to exert that pressure on him despite his request for counsel, so long as they are careful not to punctuate their statements with question marks. And if, contrary to all reasonable expectations, the suspect makes an incriminating statement, that statement can be used against him at trial. The Court thus turns *Miranda*'s unequivocal rule against any interrogation at all into a trap in which unwary suspects may be caught by police deception.

2

. . . In any event, I think the Court is clearly wrong in holding, as a matter of law, that Officer Gleckman should not have realized that his statement was likely to elicit an incriminating response. The Court implicitly assumes that, at least in the absence of a lengthy harangue, a criminal suspect will not be likely to respond to indirect appeals to his humanitarian impulses. It then goes on to state that the officers in this case had no reason to believe that respondent would be unusually susceptible to such appeals. . . . Finally, although the significance of the officer's intentions is not clear under its objective test, the Court states in a footnote that the record "in no way suggests" that Officer Gleckman's remarks were designed to elicit a response. . . .

The Court's assumption that criminal suspects are not susceptible to appeals to conscience is directly contrary to the teachings of police interrogation manuals, which recommend appealing to a suspect's sense of morality as a standard and often successful interrogation technique. Surely the practical experience embodied in such manuals should not be ignored in a case such as this in which the record is devoid of any evidence—one way or the other—as to the susceptibility of suspects in general or of Innis in particular.

NOTES

1. Any knowledge the police may have had concerning the unusual susceptibility of a defendant to a particular form of persuasion might be an important factor in determining whether the police should have known that their words or actions were reasonably likely to elicit an incriminating response from the suspect.
2. The record in no way suggests that the officers' remarks were *designed* to elicit a response. . . . It is significant that the trial judge, after hearing the officers' testimony, concluded that it was "entirely understandable that [the officers] would voice their concern [for the safety of the handicapped children] to each other."
3. In limiting its test to police statements "likely to elicit an incriminating response," the Court confuses the scope of the exclusionary rule with the definition of "interrogation." Of course, any incriminating statement as defined in *Miranda* . . . must be excluded from evidence if it is the product of impermissible interrogation. But I fail to see how this rule helps in deciding whether a particular statement or tactic constitutes "interrogation." After all, *Miranda* protects a suspect in Innis' position not simply from interrogation that is likely to be successful, but from any interrogation at all.

REVIEW AND DISCUSSION QUESTIONS

1. The Fifth Amendment states that no one "shall be compelled in any criminal case to be a witness against himself." What possible moral or political justification is there for a right against self-incrimination? What, if anything, would be morally objectionable about a judicial system that required accused persons to testify?

2. In *Miranda*, the Court describes various interrogation techniques that involve trickery and psychological manipulation. What, if anything, is morally objectionable about such techniques? The Court states that its "decision is not intended to hamper the traditional function of police officers in investigating a crime," but is effective police work possible without the use of such techniques? Do the Court's examples support McCloskey's reasons for believing that many innocent people are wrongfully convicted?
3. Why is the issue in *Brewer* one of the right to counsel rather than self-incrimination?
4. Did the "Christian burial speech" violate Williams's right to counsel? Who is right about this—the majority or the dissenters? Explain.
5. In *Innis*, the Court states that "the *Miranda* safeguards come into play whenever a person in custody is subjected to either express questioning or its functional equivalent." Was the Court justified in deciding that Innis had not been subject to interrogation in either sense? Why or why not? What factors need to be considered in determining whether a suspect like Innis was or was not interrogated?
6. Consider the three statements distinguished by Stevens in his dissent. Is he correct that all three should be considered interrogation? Would you agree with Stevens that the Court's decision creates "an incentive for police to ignore a suspect's invocation of his rights in order to make continued attempts to extract information from him"?

Entrapment

United States v. Tobias

Effective law enforcement often requires undercover police work and the close involvement of undercover officers with suspected criminals. Sometimes such law enforcement strategies can entice or enable defendants to commit crimes they would not otherwise have committed. When such strategies go too far, a criminal defendant may escape conviction if he or she was "entrapped" by the government or, alternatively, if the government was so involved with the crime as to have violated due process. In this case, involving the manufacture of illegal drugs, Judge Hatchett, writing for the court majority, rules that the government's conduct was lawful, thus upholding the conviction of the defendant, Thomas Tobias. In his dissent, Judge Johnson contends that the DEA went too far: Although the government did not "entrap" Tobias, it was "overinvolved" with the crime for which he was charged.

Hatchett, Circuit Judge: We again examine the extent to which the government may become involved in a criminal enterprise without being found guilty of entrapment or having its conduct declared so outrageous as to violate a criminal defendant's due process rights under the fifth amendment. Finding the government's conduct in this case to be within lawful bounds, we affirm.

FACTS

In order to pursue undercover investigations of clandestine laboratory operators, the Drug Enforcement Administration (DEA) established a chemical supply company in a mid-western state.... The supply company operated in the same way as any legitimate chemical supply company. It had a business location and received orders via

662 F.2d 381 (1981)

telephone and mail for various chemicals which could be used in the manufacture of controlled substances.

In the April 1980 edition of *High Times Magazine,* the DEA supply company placed an advertisement offering over-the-counter sales of chemicals and laboratory equipment. On March 18, 1980, supply company received a letter from Thomas C. Tobias, the defendant, requesting "more information" and giving his name and address. A catalog was sent to Tobias in Mobile, Alabama. On March 26, 1980, Tobias telephoned the supply company and placed an order for various chemicals. On April 15, 16, 17, and 24, 1980, Tobias telephoned the supply company to check on his order and to order additional chemicals.

On April 25, 1980, Tobias again called the supply company about his order. According to Tobias's testimony, he telephoned to cancel his order because he had discovered from reading drug literature that he could not manufacture cocaine without "more knowledge . . . and a lot of equipment." Before Tobias could cancel his order, however, Special Agent Schabilion asked him what he was trying to do. Tobias admitted that he wanted to make cocaine but had encountered difficulties. Pretending to empathize with Tobias, Schabilion stated that he too found cocaine to be extremely difficult and expensive to manufacture. To this, Tobias said he was not necessarily interested in manufacturing cocaine, but "just wanted to make some money." Agent Schabilion advised Tobias that "almost anything would be cheaper and easier to manufacture than cocaine," including amphetamines. Schabilion then suggested that Tobias make Phencyclidene (PCP). Schabilion explained that making PCP was as easy as "baking a cake" and that for $500 he would send Tobias everything he needed to get set up. Stating that he might have a market for PCP in Mobile, Tobias agreed. He cancelled his order for the original chemicals and told Special Agent Schabilion to send him everything he needed to manufacture PCP.

The supply company shipped the formula and some of the chemicals needed to manufacture PCP to the DEA office in Mobile for delivery to Tobias. It is undisputed that the chemicals provided by the DEA were not difficult to obtain and could have been purchased at other chemical supply houses. After receiving the chemicals and formula from the DEA, Tobias telephoned the supply company thirteen times to discuss problems encountered in the manufacturing process and to obtain advice for overcoming them. On May 9, 1980, when DEA agents executed a search warrant on Tobias's residence in Mobile they found PCP in a liquid state.

Tobias was convicted in a non-jury trial and received sentences totaling fifteen years in prison.

I

Tobias complains that the record is devoid of any evidence indicating that he entertained the thought of making drugs prior to reading the government's advertisement for chemicals. He also argues that even after he sought to abandon his scheme to manufacture cocaine, the DEA agents suggested that he make a "cheaper and easier" drug and provided him with the necessary precursors, equipment, and know-how. Thus, he argues that he was entrapped as a matter of law.

"[W]hen entrapment is at issue, the focal point of the inquiry is on the predisposition of the defendant." Thus, a defendant who wishes to assert an entrapment defense must initially come forward with evidence "that the Government's conduct created a substantial risk that the offense would be committed by a person other than one ready to commit it." Once the defendant has carried this burden, the government must, if it is to prevail, prove beyond a reasonable doubt that the defendant was predisposed to commit the crime charged.

"A prosecution cannot be defeated merely because a government agent has provided the accused with the opportunity or facilities for the commission of the crime. It is only when the government's deception actually implants the criminal design in the mind of the defendant that the defense of entrapment comes into play."

On this record, we cannot say that Tobias was entrapped. Even assuming Tobias produced sufficient evidence to raise the issue of entrapment, we are satisfied that the government carried its burden of proving, beyond a reasonable doubt, that Tobias was predisposed to commit the charged offenses. The government's proof showed that

Tobias responded to a simple advertisement offering the over-the-counter sale of chemicals which could be purchased without any difficulty in chemical houses in Mobile, Alabama. This advertisement served only to provide one so disposed the opportunity to obtain the necessary precursors and equipment to manufacture controlled substances. Tobias seized this opportunity by writing the supply company for "more information" and telephoning the supply company on many occasions to place and check on his order. The DEA did nothing else to solicit Tobias's business. A prosecution may not be defeated because the government provides the accused with the opportunity to commit the crimes charged.

Tobias also contends that agent Schabilion's suggestion that PCP would be "cheaper and easier" to manufacture implanted the criminal design in his mind at a time when he sought to cancel his order for chemicals necessary to manufacture cocaine.

The record simply does not bear out Tobias's contention. The record shows that although Tobias sought to cancel his original order, he indicated to agent Schabilion that he was not interested in manufacturing any particular drug but was only interested in making money. At that point, agent Schabilion suggested that amphetamines, including PCP, would be "cheaper and easier" to manufacture. Tobias then indicated that there might be a market for PCP in Mobile and asked Schabilion to send him the formula, equipment, and precursors necessary to manufacture PCP. This evidence shows that Tobias was predisposed to manufacture a controlled substance, although no one drug in particular.

If law enforcement agents are precluded from discussing the particulars of how a criminal enterprise is to be conducted, the undercover work that is essential to the investigation and prosecution of drug offenses becomes impossible. Suggestions regarding the particulars of manufacturing one drug or another did not vitiate the predisposition which is best shown by Tobias's continuance of the conversation.

II

Tobias next argues that if the government had not provided him with the formula, necessary precursors, and continuing advice during the manufacturing process, he would have been unable to manufacture PCP. Thus, he argues that the government's involvement in this scheme was so outrageous that due process principles bar his convictions. This presents a tougher question.

In *Russell,* the Supreme Court held that the defense of entrapment was foreclosed to one who was predisposed to commit a crime, regardless of the type and degree of government activity involved. The Court, however, expressed the possibility that due process principles might prohibit an excessive degree of government involvement, stating: "[W]e may some day be presented with a situation in which the conduct of law enforcement agents is so outrageous that due process principles would absolutely bar the government from invoking judicial processes to obtain a conviction. . . ." While the Court pointed out that in certain situations the conduct of the government may bar prosecution, it emphasized that the defendant must show that the challenged government conduct violates " 'that fundamental fairness, shocking to the universal sense of justice,' mandated by the due process clause of the fifth amendment."

In certain circumstances, the government's conduct may be so outrageous as to violate due process, however "[p]olice overinvolvement in crime would have to reach a demonstrable level of outrageousness before it could bar convictions."

In the recent case of *United States v. Gray,* this court was presented with the question whether the government's conduct was so outrageous as to require reversal of convictions for violation of due process. In *Gray,* two government agents suggested a smuggling scheme to defendants and provided them with repair services, an airstrip, and a crew. The court, although acknowledging that the government agents suggested the scheme and aided in arranging the air transportation, held that "the providing of essential services is not misconduct."

In *United States v. Leja,* the defendant approached a government informant and suggested that the two produce PCP in a laboratory. They entered into an agreement which provided that the government informant would obtain the necessary chemicals and the other two defendants would supply glassware, money, and technical expertise. Not only did the government informant provide the

chemicals necessary to produce PCP, but another government informant provided technical instructions concerning the manufacturing process when one of the defendants -encountered difficulties. The Sixth Circuit, relying on the fact that the government agent neither solicited the defendants nor provided them with chemicals and information which they could not have obtained elsewhere, held that the due process clause did not require reversal of their convictions.

In *United States v. Twigg,* however, the Third Circuit found the government's conduct so outrageous as to violate due process because the government agent suggested the establishment of a drug laboratory, provided the place, equipment, supplies and know-how, and then ran the entire operation with only meager assistance from the defendants.

From these cases emerges the basic proposition that government infiltration of criminal activity is a "recognized and permissible means of investigation." "This proposition remains true even though the . . . government agent . . . supplies something of value to the criminal." This is necessary so that the agent "will . . . be taken into the confidence of the illegal entrepreneurs." On the other end of the spectrum, however, the government may not instigate the criminal activity, provide the place, equipment, supplies and know-how, and run the entire operation with only meager assistance from the defendants without violating fundamental fairness.

Although the DEA provided the formula and some of the chemicals for the manufacture of PCP, the chemicals were not difficult to obtain and, in fact, some were ordered from a chemical supply house in Mobile. DEA provided no financial aid for Tobias's operation. In sum, "[t]he law enforcement conduct here stops far short of violating that 'fundamental fairness, shocking to the universal sense of justice,' mandated by the due process clause of the fifth amendment."

The cases demonstrate that outrageous involvement turns upon the totality of the circumstances with no single factor controlling. Although a totality of the circumstances standard must be applied, it is beneficial to review the parts that make up the whole. The DEA, in this case, did not initiate contact with Tobias. May the government be held to have involved itself in outrageous conduct by placing the ad in *High Times?* Similarly, may the government be condemned for shipping the necessary chemicals, even at cut-rate prices? Or, was it outrageous for DEA to deliver the chemicals to Tobias's home? We think not. The crucial factor in this total fact picture is the step-by-step advice given by the DEA agents. This advice was given to Tobias or his wife on more than thirteen occasions. On each occasion, however, Tobias or his wife contacted the DEA. This would be a more difficult case if the DEA had pursued Tobias by repeated phone calls and encouragement. But here, the drug transaction would have stopped at any time that Tobias made no further calls. Instead of being a *predisposed inactive participant* in this scheme to manufacture and distribute PCP (a cheap mind-bending drug primarily sold to youngsters), Tobias was a *predisposed active participant* motivated solely by a desire to make money. It is this predisposition plus active and insistent participation that sets this case apart from cases finding a due process violation. We are mindful of the Supreme Court's admonition that due process can only be invoked in the rarest and most outrageous circumstances. Yet, this case does set the outer limits to which the government may go in the quest to ferret out and prosecute crimes in this circuit.

Frank M. Johnson, Jr., Circuit Judge, Dissenting: Because I conclude that the Drug Enforcement Agency's (DEA) "overinvolvement" in the commission of the crime for which defendant was convicted precludes prosecution, I am unable to agree with the majority's opinion and therefore respectfully dissent.

Tobias is a 21- or 22-year-old married man employed as a house painter. With the exception of a minor exchange of fisticuffs, he has no prior record of any kind and had never been convicted of a drug related offense. In early 1980 Tobias responded to an advertisement in *High Times,* an over-the-counter drug culture periodical, and sought information concerning the purchase of chemicals necessary to make cocaine. Defendant requested a catalogue from the chemical supply company and eventually ordered what he thought were chemicals needed to manufacture cocaine. Unbeknownst to Tobias, the chemical supply company, located in a mid-western state, was owned

and operated by the DEA. On April 25, 1980, Tobias called the company to cancel his order and spoke to DEA Agent Schabilion. He advised Schabilion that he intended to use the chemicals to manufacture cocaine but had concluded that he lacked the requisite knowledge and equipment necessary to make the drug. Indeed, the record reflects that Tobias was ineptly suited to make cocaine. His education extended only as far as high school, whereas manufacturing cocaine requires a "sophisticated chemical background." Further, it appears that Tobias never had even a basic understanding of what was needed to manufacture the drug, both in terms of the necessary chemicals and in terms of the necessary knowledge of chemistry.[1] During the course of the conversation, Schabilion emphasized the difficulty and expense involved in manufacturing cocaine and suggested a number of "cheaper" and "easier" alternatives. "I advised him [Tobias] that almost anything would be cheaper and easier to manufacture than cocaine, and suggested that any of the amphetamines, speed or PCP would be easier." Until that statement, nothing in the record reveals any intent or disposition on the part of Tobias to attempt the manufacture of any drug other than cocaine. Tobias admitted to Schabilion that there might be a market for PCP in the Mobile area but expressed reservations about making the drug, stating that he had no background in chemistry. At that point in the conversation Agent Schabilion explained that manufacturing PCP was as easy as "baking a cake." The agent, according to defendant, also told him that selling the drug would be profitable. Without any discussion as to particular chemicals or quantities, Schabilion promised to supply Tobias with everything necessary to make PCP, including the formula, at a total cost of $500. Tobias then agreed to purchase the chemicals and delivery was made in Mobile, Alabama, by a DEA agent acting as an employee of a fictitious delivery company. The defendant received two separate deliveries on the same day and paid $101 for the chemicals, the total amount requested by the DEA. The DEA supplied Tobias with all of the necessary ingredients of PCP except diethyl ether, which Tobias obtained at a local chemical supply store.

Utilizing the chemicals provided by the DEA and the formula provided by the DEA, Tobias attempted to manufacture a drug suggested by the DEA. However, the agency's encouragement, pervasive influence, and active participation did not end there. Still uncertain about how to actually make the drug, from April 29 through May 9 Tobias telephoned the supply company 13 times and his wife called on three other occasions to obtain information and additional advice concerning the procedure at each stage of the manufacturing process. It is undisputed that Tobias did not know how to make PCP and that during those phone calls the DEA agents instructed him as to each step in the process. Agents for the Government, therefore, guided the defendant in every stage of the process, from inception to termination. Tobias was ultimately arrested and, despite his status as a first-time drug offender, his youthful age and his family situation, was cumulatively sentenced to fifteen years' incarceration and ten years' special parole.

Government involvement and assistance in the criminal activities of others are sometimes necessary to enforce the laws. The degree of Government involvement is not, however, boundless. The Supreme Court observed that, when involvement by federal agents in the commission of a crime is so outrageous as to shock the conscience, either due process or the Court's supervisory powers will preclude prosecution. These decisions make clear that as a general rule infiltration and "limited involvement" by Government agents in a drug related enterprise do not run afoul of the due process clause and in fact constitute a legitimate method of apprehending offenders. . . . Only extreme and outrageous Government involvement in the commission of a crime will justify reversing for that reason a conviction under the due process clause.

Although providing a rudimentary framework for analysis, neither *Russell* nor *Hampton* delineated with clarity the point at which Government involvement becomes shocking and outrageous. However, some lower court decisions have found the requisite level of outrageous involvement and provide useful guidance. In *Greene v. United States,* a case cited by the Supreme Court in *Hampton,* the defendants contacted a Government agent concerning the possibility of manufacturing and selling boot-leg whiskey. The agent represented himself to be a member of the "syndicate" and agreed to

both provide financial assistance and buy all of the whiskey produced. Extensive communications transpired between the parties, with the agent providing the defendants with substantial assistance. He helped search for an appropriate site for the still, agreed to provide equipment and an operator for the still, and made available two thousand pounds of sugar at wholesale prices. The agent also tried to prompt the defendants by intimating that he was receiving pressure for the whiskey from his syndicate boss. On appeal, the court noted that the defendants evinced a predisposition to commit the crime and could not, therefore, invoke the "usual entrapment defense." However, because of the extensive and aggressive overinvolvement by the agent in the crime, the conviction was reversed. "[W]hen the Government permits itself to become enmeshed in criminal activity, from beginning to end, to the extent which appears here, the same underlying objections which render entrapment repugnant to American criminal justice are operative."

Similarly in *United States v. Twigg,* AED agents initiated contact with the defendants and suggested that they construct a laboratory for manufacturing amphetamines. The DEA agents supplied the defendants with the necessary chemicals at a discount price, supplied a portion of glassware to be used, and provided a barn as the situs for the laboratory. The agents facilitated defendants' ability to purchase chemicals from other supply companies and made available the necessary funds to make the purchases. A DEA agent also supervised the entire manufacturing process, with the defendants providing only minor production assistance. The Third Circuit rejected the defense of entrapment, finding evidence of predisposition. However, the court reversed the conviction and held that "the nature and extent of police involvement in this crime was so overreaching as to bar prosecution of the defendants as a matter of law."

The cases demonstrate that outrageous involvement turns upon the totality of the circumstances with no single factor controlling. Under such a standard, it is my judgment that the DEA's overinvolvement in this case requires reversal of the conviction. The DEA agents placed the advertisement for the chemicals in *High Times* magazine. The majority justifies this conduct by observing that it was for the purpose of pursuing "undercover investigations of clandestine laboratory operators. . . ." While that may have been the purpose of the operation of the chemical supply house, in this case the advertisement constituted the initial contact with a young man that had up until then indicated no predisposition to manufacture drugs. When it became apparent that Tobias was completely stymied in his predisposition to manufacture cocaine, the agents advanced the idea of manufacturing an "easier" and "cheaper" drug. Tobias was provided with almost all of the necessary chemicals at apparently cut-rate prices, was provided with the formula for manufacturing the PCP and was persuaded to make PCP in part by the statement that the process was no more difficult than "baking a cake." The DEA agents personally delivered the chemicals to his home and gave technical advice concerning the amount and sequence of combining each separate ingredient at every step of the manufacturing process. Admittedly, the agents did not other than the advertisement initiate contact with Tobias, but that fact is not dispositive.[2]

This case does not merely involve law enforcement officers supplying Tobias with an "item of value" necessary to make the prohibited drug. Such a practice is universally recognized as an acceptable investigative technique. Here, Tobias amounted to little more than a conduit. Had the Government agents at any point ceased providing Tobias with assistance and encouragement, the record indicates that he would have been incapable of manufacturing the illicit drug. Thus Government agents vicariously manufactured PCP through Tobias in order to gather evidence to justify his prosecution. The instant case therefore involves that rare degree of Government involvement in the commission of a crime that *Hampton* and *Russell* intended to prevent. I do not, of course, propose that federal courts be accorded a "chancellor's foot" veto over law enforcement techniques that are disapproved.

I do suggest, however, that the majority goes too far by holding, in effect, that no matter how egregious and shocking the degree of Government involvement in the commission of a crime, neither the due process clause nor the court's supervisory powers will ever be invoked to preclude prosecution.

Courts are rightfully loath to overturn convictions on the basis of Governmental overinvolvement in the commission of a crime, recognizing that law enforcement officials need wide latitude in devising appropriate methods and tactics for apprehending violators. However, judicial tolerance of Governmental involvement in criminal activity should not constitute a *carte blanche;* there is still a point beyond which law enforcement officials cannot go. As Justice Brandeis observed, "[t]o declare that in the administration of criminal law the end justifies the means—to declare that the Government may commit crimes in order to secure the conviction of a private criminal—would bring terrible retribution." The direct, continual involvement by Government agents in the creation, maintenance and commission of a crime, to the extent reflected in this case, goes beyond the perimeter of permissible conduct and should not be countenanced by this Court.

NOTES

1. Tobias originally ordered four chemicals from the supply company and later added two more. According to counsel for the Government, DEA officials did not ship the chemicals because they were unable to tell from the order what Tobias was trying to manufacture. Apparently Tobias' list of chemicals was either incomplete or inaccurate, indicating that he did not know the proper ingredients. Further, it was Agent Schabilion during the April 25 phone call who had to explain to Tobias about the chemicals, precursors, expense and technical knowledge necessary to make cocaine.
2. The majority indicates that the DEA's failure to provide Tobias with "financial aid" distinguishes the case from prior decisions dealing with Government overinvolvement. The record indicates, however, that the Government did provide Tobias with "financial assistance." Tobias was originally informed that the chemicals would cost $500. Nonetheless, when the DEA agent delivered the chemicals, he charged Tobias only $101. Moreover, the DEA agent made a second delivery on the same day and did not charge Tobias *anything* for the additional chemicals. Thus, although the record is not totally clear, the Government appears to have provided "financial assistance" to Tobias in the form of reduced prices for the chemicals used to manufacture the PCP.

REVIEW AND DISCUSSION QUESTIONS

1. What are the facts of this case? How, if at all, does the case differ from *Greene* and *Twigg* (discussed by Judge Johnson)?
2. What are the main arguments that Tobias raised in his defense, and how does Judge Hatchett respond to them? Are you persuaded by Hatchett's reasoning?
3. How is "predisposition" relevant to the entrapment defense? Was Tobias in fact disposed to commit the crime for which he was charged? Judge Hatchett distinguishes between a "predisposed inactive participant" and a "predisposed active participant"? In your view, was Tobias an active or inactive participant?
4. Why does Judge Johnson believe that the DEA was "overinvolved" with the crime? With whom do you agree about this, Judge Hatchett or Judge Johnson?
5. Why should a legal system permit an entrapment defense in the first place? How should entrapment be defined? In *Greene,* the appellate court held that a conviction could be reversed because government involvement with the crime was too extensive and aggressive, even though the defendant was not entrapped. Do you agree with this idea? Explain why or why not.

CHAPTER 14

Sentencing

Philosophers of law have paid surprising little attention to the sentencing of criminals, even though the subject has clear connections to other areas of philosophical interest in the criminal law such as the justification of punishment, the rights of defendants, and the nature of criminal responsibility. Moreover, the question of the specific types of punishment and the particular sentencing practices that a legal system should employ if it is to be both efficient and just is of the highest moral and philosophical importance. The readings in this section explore this topic. In *Gregg v. Georgia,* the Supreme Court airs arguments for and against the death penalty, while in "Electric Shock: The Fairest Punishment of All," Graeme Newman makes the case for reviving corporal punishment as a basic component of our criminal justice system. In *Rummell v. Estelle,* the Supreme Court examines whether recidivist or "three strikes and you're out" statutes sometimes impose a punishment that is unconstitutional because it is "cruel and unusual." In *McClesky v. Kemp,* the Court takes up the legal implication of racial disparities in capital punishment cases.

Capital Punishment

Gregg v. Georgia

One of the most controversial issues in sentencing concerns the justifiability of the death penalty. In *Furman v. Georgia* (1974) the Supreme Court ruled that in Georgia (and by implication most other states) whether a convicted murderer received the death penalty or imprisonment was arbitrary and capricious and, hence, unconstitutional. In response, Georgia and many other states revised their procedures for determining when and under what circumstances a sentence of death was to be handed down, basically making this decision more uniform and objective. With Georgia and other states having now addressed

428 U.S. 153 (1976). Some footnotes and citations omitted.

the Supreme Court's concerns about the unconstitutional administration of the death penalty, the question before the Court in *Gregg v. Georgia* is whether capital punishment is unconstitutional even when administered fairly; that is, whether in principle capital punishment violates the Eighth Amendment's prohibition of "cruel and unusual punishment." *Greg* is one of the U.S. Supreme Court's most important decisions. As is apparent from the dissenting opinions, few of the issues it addresses involve deeper legal, moral, and philosophical division than does capital punishment.

Mr. Justice Stewart, with Justices Powell and Stevens concurring: We address initially the basic contention that the punishment of death for the crime of murder is, under all circumstances, "cruel and unusual" in violation of the Eighth and Fourteenth Amendments of the Constitution. [Later in] this opinion, we will consider the sentence of death imposed under the Georgia statutes at issue in this case. . . .

The substantive limits imposed by the Eighth Amendment on what can be made criminal and punished were discussed in *Robinson v. California,* 370 U.S. 660 (1962). The Court found unconstitutional a state statute that made the status of being addicted to a narcotic drug a criminal offense. It held, in effect, that it is "cruel and unusual" to impose any punishment at all for the mere status of addiction. The cruelty in the abstract of the actual sentence imposed was irrelevant: "Even one day in prison would be a cruel and unusual punishment for the 'crime' of having a common cold." Most recently, in *Furman v. Georgia* . . . three Justices in separate concurring opinions found the Eighth Amendment applicable to procedures employed to select convicted defendants for the sentence of death.

It is clear from the foregoing precedents that the Eighth Amendment has not been regarded as a static concept. As Mr. Chief Justice Warren said, in an oft-quoted phrase, "[t]he Amendment must draw its meaning from the evolving standards of decency that mark the progress of a maturing society." . . . Thus, an assessment of contemporary values concerning the infliction of a challenged sanction is relevant to the application of the Eighth Amendment. As we develop below more fully, . . . this assessment does not call for a subjective judgment. It requires, rather, that we look to objective indicia that reflect the public attitude toward a given sanction.

But our cases also make clear that public perceptions of standards of decency with respect to criminal sanctions are not conclusive. A penalty also must accord with "the dignity of man," which is the "basic concept underlying the Eighth Amendment." *Trop v. Dulles.* . . . This means, at least, that the punishment not be "excessive." When a form of punishment in the abstract (in this case, whether capital punishment may ever be imposed as a sanction for murder) rather than in the particular (the propriety of death as a penalty to be applied to a specific defendant for a specific crime) is under consideration, the inquiry into "excessiveness" has two aspects. First, the punishment must not involve the unnecessary and wanton infliction of pain. . . . Second, the punishment must not be grossly out of proportion to the severity of the crime. . . .

The imposition of the death penalty for the crime of murder has a long history of acceptance both in the United States and in England. The common-law rule imposed a mandatory death sentence on all convicted murderers. . . . And the penalty continued to be used into the 20th century by most American States, although the breadth of the common-law rule was diminished, initially by narrowing the class of murders to be punished by death and subsequently by widespread adoption of laws expressly granting juries the discretion to recommend mercy. . . .

It is apparent from the text of the Constitution itself that the existence of capital punishment was accepted by the Framers. At the time the Eighth Amendment was ratified, capital punishment was a common sanction in every State. Indeed, the First Congress of the United States enacted legislation providing death as the penalty for specified crimes. . . .

The most marked indication of society's endorsement of the death penalty for murder is the legislative response to *Furman.* The legislatures of at least 35 States have enacted new statutes that provide for the death penalty for at least some

crimes that result in the death of another person. And the Congress of the United States, in 1974, enacted a statute providing the death penalty for aircraft piracy that results in death. These recently adopted statutes have attempted to address the concerns expressed by the Court in *Furman* primarily (i) by specifying the factors to be weighed and the procedures to be followed in deciding when to impose a capital sentence, or (ii) by making the death penalty mandatory for specified crimes.

... [H]owever, the Eighth Amendment demands more than that a challenged punishment be acceptable to contemporary society. The Court also must ask whether it comports with the basic concept of human dignity at the core of the Amendment.... Although we cannot "invalidate a category of penalties because we deem less severe penalties adequate to serve the ends of penology," *Furman v. Georgia, supra,* at 451 (Powell, J., dissenting), the sanction imposed cannot be so totally without penological justification that it results in the gratuitous infliction of suffering....

The death penalty is said to serve two principal social purposes: retribution and deterrence of capital crimes by prospective offenders.[1]

In part, capital punishment is an expression of society's moral outrage at particularly offensive conduct. This function may be unappealing to many, but it is essential in an ordered society that asks its citizens to rely on legal processes rather than self-help to vindicate their wrongs.

> The instinct for retribution is part of the nature of man, and channeling that instinct in the administration of criminal justice serves an important purpose in promoting the stability of a society governed by law. When people begin to believe that organized society is unwilling or unable to impose upon criminal offenders the punishment they "deserve," then there are sown the seeds of anarchy—of self-help, vigilante justice, and lynch law. *Furman v. Georgia*, supra, at 308 (Stewart, J., concurring).

"Retribution is no longer the dominant objective of the criminal law," *Williams v. New York,* 337 U.S. 241, 248 (1949), but neither is it a forbidden objective nor one inconsistent with our respect for the dignity of men.... Indeed, the decision that capital punishment may be the appropriate sanction in extreme cases is an expression of the community's belief that certain crimes are themselves so grievous an affront to humanity that the only adequate response may be the penalty of death.

Statistical attempts to evaluate the worth of the death penalty as a deterrent to crimes by potential offenders have occasioned a great deal of debate. The results simply have been inconclusive....

Although some of the studies suggest that the death penalty may not function as a significantly greater deterrent than lesser penalties, there is no convincing empirical evidence either supporting or refuting this view. We may nevertheless assume safely that there are murderers, such as those who act in passion, for whom the threat of death has little or no deterrent effect. But for many others, the death penalty undoubtedly is a significant deterrent. There are carefully contemplated murders, such as murders for hire, where the possible penalty of death may well enter into the cold calculus that precedes the decision to act.[2] And there are some categories of murder, such as murder by a life prisoner, where other sanctions may not be adequate.

The value of capital punishment as a deterrent of crime is a complex factual issue the resolution of which properly rests with the legislatures, which can evaluate the results of statistical studies in terms of their own local conditions and with a flexibility of approach that is not available to the courts.... Indeed, many of the post-*Furman* statutes reflect just such a responsible effort to define those crimes and those criminals for which capital punishment is most probably an effective deterrent.

In sum, we cannot say that the judgment of the Georgia Legislature that capital punishment may be necessary in some cases is clearly wrong. Considerations of federalism, as well as respect for the ability of a legislature to evaluate, in terms of its particular State, the moral consensus concerning the death penalty and its social utility as a sanction, require us to conclude, in the absence of more convincing evidence, that the infliction of death as a punishment for murder is not without justification and thus is not unconstitutionally severe.

Finally, we must consider whether the punishment of death is disproportionate in relation to the crime for which it is imposed. There is no question that death as a punishment is unique in its severity and irrevocability.... When a defendant's life is at

stake, the Court has been particularly sensitive to insure that every safeguard is observed. . . . But we are concerned here only with the imposition of capital punishment for the crime of murder, and when a life has been taken deliberately by the offender, we cannot say that the punishment is invariably disproportionate to the crime. It is an extreme sanction, suitable to the most extreme of crimes.

We hold that the death penalty is not a form of punishment that may never be imposed, regardless of the circumstances of the offense, regardless of the character of the offender, and regardless of the procedure followed in reaching the decision to impose it.

We now consider whether Georgia may impose the death penalty on the petitioner in this case. . . .

Because of the uniqueness of the death penalty, *Furman* held that it could not be imposed under sentencing procedures that created a substantial risk that it would be inflicted in an arbitrary and capricious manner. Mr. Justice White concluded that "the death penalty is exacted with great infrequency even for the most atrocious crimes and . . . there is no meaningful basis for distinguishing the few cases in which it is imposed from the many cases in which it is not." 408 U.S., at 313 (concurring). Indeed, the death sentences examined by the Court in *Furman* were "cruel and unusual in the same way that being struck by lightning is cruel and unusual. For, of all the people convicted of [capital crimes], many just as reprehensible as these, the petitioners [in *Furman* were] among a capriciously selected random handful upon whom the sentence of death has in fact been imposed. . . . [T]he Eighth and Fourteenth Amendments cannot tolerate the infliction of a sentence of death under legal systems that permit this unique penalty to be so wantonly and so freakishly imposed." *Id.*, at 309–310 (Stewart, J., concurring). . . .

In summary, the concerns expressed in *Furman* that the penalty of death not be imposed in an arbitrary or capricious manner can be met by a carefully drafted statute that ensures that the sentencing authority is given adequate information and guidance. As a general proposition these concerns are best met by a system that provides for a bifurcated proceeding at which the sentencing authority is apprised of the information relevant to the imposition of sentence and provided with standards to guide its use of the information. . . .

The basic concern of *Furman* centered on those defendants who were being condemned to death capriciously and arbitrarily. Under the procedures before the Court in that case, sentencing authorities were not directed to give attention to the nature or circumstances of the crime committed or to the character or record of the defendant. Left unguided, juries imposed the death sentence in a way that could only be called freakish. The new Georgia sentencing procedures, by contrast, focus the jury's attention on the particularized nature of the crime and the particularized characteristics of the individual defendant. While the jury is permitted to consider any aggravating or mitigating circumstances, it must find and identify at least one statutory aggravating factor before it may impose a penalty of death. In this way the jury's discretion is channeled. No longer can a jury wantonly and freakishly impose the death sentence; it is always circumscribed by the legislative guidelines. In addition, the review function of the Supreme Court of Georgia affords additional assurance that the concerns that prompted our decision in *Furman* are not present to any significant degree in the Georgia procedure applied here.

For the reasons expressed in this opinion, we hold that the statutory system under which Gregg was sentenced to death does not violate the Constitution. Accordingly, the judgment of the Georgia Supreme Court is affirmed.

It is so ordered.

Mr. Justice Brennan, dissenting: The Cruel and Unusual Punishments Clause "must draw its meaning from the evolving standards of decency that mark the progress of a maturing society."[3] The opinions of Mr. Justice Stewart, Mr. Justice Powell, and Mr. Justice Stevens today hold that "evolving standards of decency" require focus not on the essence of the death penalty itself but primarily upon the procedures employed by the State to single out persons to suffer the penalty of death. Those opinions hold further that, so viewed, the Clause invalidates the mandatory infliction of the death penalty but not its infliction under sentencing procedures that Mr. Justice Stewart, Mr. Justice Powell,

and Mr. Justice Stevens conclude adequately safeguard against the risk that the death penalty was imposed in an arbitrary and capricious manner.

In *Furman v. Georgia,* . . . I read "evolving standards of decency" as requiring focus upon the essence of the death penalty itself and not primarily or solely upon the procedures under which the determination to inflict the penalty upon a particular person was made. I there said:

> From the beginning of our Nation, the punishment of death has stirred acute public controversy. Although pragmatic arguments for and against the punishment have been frequently advanced, this longstanding and heated controversy cannot be explained solely as the result of differences over the practical wisdom of a particular government policy. At bottom, the battle has been waged on moral grounds. The country has debated whether a society for which the dignity of the individual is the supreme value can, without a fundamental inconsistency, follow the practice of deliberately putting some of its members to death. In the United States, as in other nations of the western world, "the struggle about this punishment has been one between ancient and deeply rooted beliefs in retribution, atonement or vengeance on the one hand, and, on the other, beliefs in the personal value and dignity of the common man that were born of the democratic movement of the eighteenth century, as well as beliefs in the scientific approach to an understanding of the motive forces of human conduct, which are the result of the growth of the sciences of behavior during the nineteenth and twentieth centuries." It is this essentially moral conflict that forms the backdrop for the past changes in and the present operation of our system of imposing death as a punishment for crime. *Id.,* at 296.[4]

That continues to be my view. For the Clause forbidding cruel and unusual punishments under our constitutional system of government embodies in unique degree moral principles restraining the punishments that our civilized society may impose on those persons who transgress its laws. Thus, I too say: "For myself, I do not hesitate to assert the proposition that the only way the law has progressed from the days of the rack, the screw and the wheel is the development of moral concepts, or, as stated by the Supreme Court . . . the application of 'evolving standards of decency.' "[5] . . .

This Court inescapably has the duty, as the ultimate arbiter of the meaning of our Constitution, to say whether, when individuals condemned to death stand before our Bar, "moral concepts" require us to hold that the law has progressed to the point where we should declare that the punishment of death, like punishments on the rack, the screw, and the wheel, is no longer morally tolerable in our civilized society. . . . I emphasize only that foremost among the "moral concepts" recognized in our cases and inherent in the Clause is the primary moral principle that the State, even as it punishes, must treat its citizens in a manner consistent with their intrinsic worth as human beings—a punishment must not be so severe as to be degrading to human dignity. A judicial determination whether the punishment of death comports with human dignity is therefore not only permitted but compelled by the Clause. . . .

The fatal constitutional infirmity in the punishment of death is that it treats "members of the human race as nonhumans, as objects to be toyed with and discarded. [It is] thus inconsistent with the fundamental premise of the Clause that even the vilest criminal remains a human being possessed of common human dignity." *Id.,* at 273. As such it is a penalty that "subjects the individual to a fate forbidden by the principle of civilized treatment guaranteed by the [clause]." I therefore would hold, on that ground alone, that death is today a cruel and unusual punishment prohibited by the Clause. "Justice of this kind is obviously no less shocking than the crime itself, and the new 'official' murder, far from offering redress for the offense committed against society, adds instead a second defilement to the first."[6]

Mr. Justice Marshall, dissenting: In *Furman* I concluded that the death penalty is constitutionally invalid for two reasons. First, the death penalty is excessive. And second, the American people, fully informed as to the purposes of the death penalty and its liabilities, in my view reject it as morally unacceptable.

Since the decision in *Furman,* the legislatures of 35 States have enacted new statutes authorizing the imposition of the death sentence for certain crimes, and Congress has enacted a law providing the death penalty for air piracy resulting in death. . . . I would be less than candid if I did not acknowledge that

these developments have a significant bearing on a realistic assessment of the moral acceptability of the death penalty to the American people. But if the constitutionality of the death penalty turns, as I have urged, on the opinion of an *informed* citizenry, then even the enactment of new death statutes cannot be viewed as conclusive. In *Furman,* I observed that the American people are largely unaware of the information critical to a judgment on the morality of the death penalty, and concluded that if they were better informed they would consider it shocking, unjust, and unacceptable. A recent study, conducted after the enactment of the post-*Furman* statutes, has confirmed that the American people know little about the death penalty, and that the opinions of an informed public would differ significantly from those of a public unaware of the consequences and effects of the death penalty.[7]

Even assuming, however, that the post-*Furman* enactment of statutes authorizing the death penalty renders the prediction of the views of an informed citizenry an uncertain basis for a constitutional decision, the enactment of those statutes has no bearing whatsoever on the conclusion that the death penalty is unconstitutional because it is excessive. An excessive penalty is invalid under the Cruel and Unusual Punishments Clause "even though popular sentiment may favor" it. . . . The inquiry here, then, is simply whether the death penalty is necessary to accomplish the legitimate legislative purposes in punishment, or whether a less severe penalty—life imprisonment—would do as well. . . .

The two purposes that sustain the death penalty as nonexcessive in the Court's view are general deterrence and retribution. In *Furman,* I canvassed the relevant data on the deterrent effect of capital punishment. . . . The state of knowledge at that point, after literally centuries of debate, was summarized as follows by a United Nations Committee:

> It is generally agreed between the retentionists and abolitionists, whatever their opinions about the validity of comparative studies of deterrence, that the data which now exist show no correlation between the existence of capital punishment and lower rates of capital crime.[8]

The available evidence, I concluded in *Furman,* was convincing that "capital punishment is not necessary as a deterrent to crime in our society." . . .

The other principal purpose said to be served by the death penalty is retribution. . . . It is this notion that I find to be the most disturbing aspect of today's unfortunate decisions.

The concept of retribution is a multifaceted one, and any discussion of its role in the criminal law must be undertaken with caution. On one level, it can be said that the notion of retribution or reprobation is the basis of our insistence that only those who have broken the law be punished, and in this sense the notion is quite obviously central to a just system of criminal sanctions. But our recognition that retribution plays a crucial role in determining who may be punished by no means requires approval of retribution as a general justification for punishment.[9] It is the question whether retribution can provide a moral justification for punishment—in particular, capital punishment—that we must consider.

My Brothers Stewart, Powell, and Stevens offer the following explanation of the retributive justification for capital punishment:

> The instinct for retribution is part of the nature of man, and channeling that instinct in the administration of criminal justice serves an important purpose in promoting the stability of a society governed by law. When people begin to believe that organized society is unwilling or unable to impose upon criminal offenders the punishment they "deserve," then there are sown the seeds of anarchy—of self-help, vigilante justice, and lynch law.

This statement is wholly inadequate to justify the death penalty. As my Brother Brennan stated in *Furman,* "[t]here is no evidence whatever that utilization of imprisonment rather than death encourages private blood feuds and other disorders." It simply defies belief to suggest that the death penalty is necessary to prevent the American people from taking the law into their own hands.

In a related vein, it may be suggested that the expression of moral outrage through the imposition of the death penalty serves to reinforce basic moral values—that it marks some crimes as particularly offensive and therefore to be avoided. The argument is akin to a deterrence argument, but differs in that it contemplates the individual's shrinking from antisocial conduct, not because he fears punishment, but because he has been told in

the strongest possible way that the conduct is wrong. This contention, like the previous one, provides no support for the death penalty. It is inconceivable that any individual concerned about conforming his conduct to what society says is "right" would fail to realize that murder is "wrong" if the penalty were simply life imprisonment.

The foregoing contentions—that society's expression of moral outrage through the imposition of the death penalty pre-empts the citizenry from taking the law into its own hands and reinforces moral values—are not retributive in the purest sense. They are essentially utilitarian in that they portray the death penalty as valuable because of its beneficial results. These justifications for the death penalty are inadequate because the penalty is, quite clearly I think, not necessary to the accomplishment of those results.

There remains for consideration, however, what might be termed the purely retributive justification for the death penalty—that the death penalty is appropriate, not because of its beneficial effect on society, but because the taking of the murderer's life is itself morally good. Some of the language of the opinion of my Brothers Stewart, Powell, and Stevens in No. 74–6257 appears positively to embrace this notion of retribution for its own sake as a justification for capital punishment. They state:

> [T]he decision that capital punishment may be the appropriate sanction in extreme cases is an expression of the community's belief that certain crimes are themselves so grievous an affront to humanity that the only adequate response may be the penalty of death.

They then quote with approval from Lord Justice Denning's remarks before the British Royal Commission on Capital Punishment:

> The truth is that some crimes are so outrageous that society insists on adequate punishment, because the wrong-doer deserves it, irrespective of whether it is a deterrent or not.

Of course, it may be that these statements are intended as no more than observations as to the popular demands that it is thought must be responded to in order to prevent anarchy. But the implication of the statements appears to me to be quite different—namely, that society's judgment that the murderer "deserves" death must be respected not simply because the preservation of order requires it, but because it is appropriate that society make the judgment and carry it out. It is this latter notion, in particular, that I consider to be fundamentally at odds with the Eighth Amendment. . . . The mere fact that the community demands the murderer's life in return for the evil he has done cannot sustain the death penalty, for as Justices Stewart, Powell, and Stevens remind us, "the Eighth Amendment demands more than that a challenged punishment be acceptable to contemporary society." . . . To be sustained under the Eighth Amendment, the death penalty must "compor[t] with the basic concept of human dignity at the core of the Amendment," ibid.; the objective in imposing it must be "[consistent] with our respect for the dignity of [other] men." . . . Under these standards, the taking of life "because the wrongdoer deserves it" surely must fall, for such a punishment has as its very basis the total denial of the wrongdoer's dignity and worth.

The death penalty, unnecessary to promote the goal of deterrence or to further any legitimate notion of retribution, is an excessive penalty forbidden by the Eighth and Fourteenth Amendments. I respectfully dissent from the Court's judgment upholding the sentences of death imposed upon the petitioners in these cases.

NOTES

1. Another purpose that has been discussed is the incapacitation of dangerous criminals and the consequent prevention of crimes that they may otherwise commit in the future.
2. Other types of calculated murders, apparently occurring with increasing frequency, include the use of bombs or other means of indiscriminate killings, the extortion murder of hostages or kidnap victims, and the execution-style killing of witnesses to a crime.
3. *Trop v. Dulles,* 356 U.S. 86, 101 (1958) (plurality opinion of Warren, C.J.).
4. Quoting T. Sellin, *The Death Penalty, A Report for the Model Penal Code Project of the American Law Institute* 15 (1959).
5. *Novak v. Beto,* 453 F.2d 661, 672 (CAS 1971) (Tuttle, J., concurring in part and dissenting in part).
6. A. Camus, *Reflections on the Guillotine,* 5–6 (Fridtjof-Karla Pub. 1960).

7. Sarat & Vidmar, "Public Opinion, The Death Penalty and the Eighth Amendment: Testing the Marshall Hypothesis." 1976 *Wis. L. Rev.* 171.
8. United Nations, Department of Economic and Social Affairs, *Capital Punishment*, pt. II, ¶ 159, p. 123 (1968).
9. See, *e.g.*, H. Hart, *Punishment and Responsibility*, 8–10, 71–83 (1968); H. Packer, *Limits of the Criminal Sanction* 38–39, 66 (1968).

REVIEW AND DISCUSSION QUESTIONS

1. According to the Court, the Eighth Amendment rules out "excessive" punishment. What two issues does the question of excessiveness raise? In your view, is capital punishment excessive in either respect?
2. The Court states that the death penalty serves two main social purposes. What are they? Can execution be morally justified either way?
3. What is the basic position of the Court, and how does its decision in this case compare with its earlier decision in the *Furman* case? On what terms is the Court now willing to permit executions?
4. What is Justice Brennan's view of the constitutionality of capital punishment? Is his reasoning persuasive?
5. Does Justice Marshall successfully rebut the argument of the majority concerning deterrence? Retribution? Whose position is more convincing—the majority's or the dissenters'?

Electric Shock: The Fairest Punishment of All

Graeme Newman

In the United States, the corporal or bodily punishment of lawbreakers is extremely rare, and many people assume that any such punishments would be "cruel and unusual" and thus in violation of the Eighth Amendment. In this essay, however, Graeme Newman argues that the use of corporal punishment—in particular, the administration of electric shocks to convicted criminals—is a fairer and more effective form of punishment than imprisonment.

The acute punishment of electric shock is easily demonstrated to be superior in every respect to our current punishment practices. Compare a typical occurrence in today's courtroom with what we would have in the future if only we could get it straight that it is pain, pure and simple, that is the essence of punishment.

The judge peers out over his glasses at the pathetic woman who sits across the courtroom. In a violent outburst she has just called him a heartless tyrant or something to that effect. The public defender and a courtroom guard restrain her.

"Mrs. Washington," says the judge. "This is your third shoplifting offense. You leave me no choice . . ."

He hesitates, expecting another outburst. Mrs. Washington's three-year-old daughter sits next to her, eyes wide and watery. The judge tries to avoid her gaze.

"Mrs. Washington, it is the judgment of this court that you be sentenced to a minimum of six months in the penitentiary and a maximum of one year. Your daughter will be turned over to the care of the Department of Youth, since the pre-sentence report indicates that you have no husband or relatives who could care adequately for her . . ."

The mother is led, crying, out of the courtroom. The child pulls at her mother's skirt, crying "Mama! Mama!" But the hands of the court are upon her, and an innocent child is about to be punished for the crime of having a guilty mother.

Every day, all across America many, many families and relatives of offenders suffer in this way. This means that literally thousands of people are punished for other people's crimes.

Now, an example of what punishment of the future could be like.

Twenty-year old John Jefferson stands beside his lawyer, the public defender.

"John Jefferson," says the judge, "the court has found you guilty of burglary in the first degree. Because this is your first offense, but the damage you did was considerable, I sentence you to . . ." The judge pushes a few buttons at his computer console. The average sentence for similar cases to Jefferson's flashes on the display, ". . . five shock units."

"You will be taken immediately to the punishment hall to receive five shock units. Next case."

The victim of this crime is sitting at the back of the court. He approaches the court clerk, who directs him to the punishment hall where he will be able to watch the administration of the punishment.

Jefferson's wife and child are ushered to the waiting room where they will await Jefferson's return after he has been punished.

Meanwhile, in the punishment hall, Jefferson is seated in a specially designed chair. As part of the arrest procedure he has already received a medical examination to establish that he was fit to receive punishment.

In addition to the victim, a few members of the press are seated on the other side of the glass screen. The punishment technician, having settled the offender in the chair, returns to an adjoining room where he can observe the offender through a one-way screen. A medic is also present.

The technician sets the machine at the appropriate pain level, turns the dial to "5," and presses the button.

Jefferson receives five painful jolts of electricity to his buttocks. He screams loudly, and by the time the punishment is over, he is almost crying with pain.

The technician returns and releases the offender. "Stand and walk a little," he says.

Jefferson walks around, rubbing his buttocks. A shade drops over the spectators' screen.

"Do you still feel the pain?" asks the medic.

"Goddam, I sure do! But it's getting better. Can I go now?"

"Just sign here, and you've paid your dues."

Jefferson sighs happily and asks, "Which way to the waiting room?"

"Straight down the passage and second left."

Jefferson enters the waiting room where his wife rushes into his arms, crying, "I'm so glad it's over! Thank goodness you weren't sent to prison."

We see in this example that only the guilty person is punished. The punishment administered is clean, simple, and most importantly, convincingly painful. It is over in a brief time, and the offender is able to return to his family and his job. Punishment is confined only to the guilty. The side effects of punishment are minimized.

There is little doubt that such punishment procedures are more than feasible. Yet . . . a major objection to it has been that, because it is subjective, it would be felt so differently by each individual that it would be an inequitable punishment. Although it is true that people do respond differently to pain, it is also true that in the area of physical pain these differences can be more easily overcome than with other kinds of pain.

Let us look at the evidence.

MEASURING DIFFERENCES IN RESPONSE TO PAIN

In general, it has been found that individuals vary according to the threshold at which they report pain (that is, the point of severity in the painful stimulus at which they report that it "hurts" or request that it be stopped). . . .

We should make an important distinction here between pain threshold and tolerance of pain.

Pain threshold refers to the point on the scale of severity of the pain stimulus (for example, heat) at which the person reports that he "feels" it as pain (in this case, prickly heat). This may, of course, vary somewhat among individuals, but it should be realized that it is usually only a painful stimulus at a very low level of intensity. As we increase the intensity of the painful stimulus the person will soon decide where along that scale to say "stop."

There is an additional advantage in the administration of physical pain in that one does not need to rely exclusively on verbal reports to ascertain when it "hurts." Rather, there are physiological reactions such as sweating and pupil dilation, which are good

indicators that the painful stimulus is indeed having a painful effect. In this way we are able to eliminate at least one aspect of the complicated process of the person's perception of pain (that is, we do not have to depend on him to tell us when it hurts).

Tolerance of pain is not necessarily related to a person's pain threshold. While a person may call out "stop" relatively early in the application of a gradually increasing amount of pain, he may nevertheless be able to stand certain levels of pain for quite some time. The tolerance of pain refers to the time element in pain. Again, people may vary in the extent to which they can stand pain over time, and it has generally been found that the variations in tolerance of pain are greater than the variations in pain threshold.

MAKING PAIN THE SAME FOR EVERYONE

Since pain in these experiments is almost always at a very low level of intensity, people will cry out or ask for it to be stopped long before it reaches a point where they can no longer stand it. Thus, we would expect variations among individuals as to the point at which they called "stop." However, if one were to administer a painful stimulus which was, say, twenty times that of the lowest pain threshold, the extent to which this pain was felt differently among individuals would be "levelled." That is, it would be felt the same (very painful) by everyone. And if we used a physiological indicator we could be even more certain.

For example, suppose we have established from our experiments that the range at which people display a high sweat reaction is from a low of 5 volts to a high of 20 volts, with most people at about the middle, that is, 12 volts. We could safely administer 30 volts for a one second duration, and be sure that everyone felt the shock, and that it was very painful. This would have the effect of levelling the punishment in that we could be absolutely sure that the punishment really hurt every person to whom it was applied, and hurt them equally.

The important point about this method of applying an amount of pain is that not only is it of brief duration, but it is certainly painful. Furthermore, the scientific studies have found that the variations in perceptions of pain are much greater for chronic pain (that is, drawn out pain, or pain tolerance) than for acute pain. We are on much safer ground using momentary application of a painful stimulus that we know will really hurt every person who receives it, than to apply some other form of pain which by its nature requires that it be administered over long periods of time. The longer the time period, the more the concept of pain tolerance will override the notion of pain threshold. And the longer the time period, the harder it is to control the amount of pain administered, and the way it is perceived.

With the application of acute pain, intensity depends much more on the amount of electric shock, the amount of heat, or amount of pressure. Only as a secondary device need time be used to vary the amount of pain. Thus, we may apply our 30 volts of shock for as brief a time as a fraction of a second. And, for some offenses, this may be sufficient.

DOES PAIN DIFFER ACCORDING TO SOCIAL AND ETHNIC BACKGROUND?

There have been a number of studies by various anthropologists and psychologists which claim to have found differential responses to pain and suffering according to religious and ethnic background. . . . In fact, all these studies are irrelevant because they measure tolerance of pain, rather than pain threshold. They simply suggest the obvious: that particular ethnic or religious groups complain differently about pain. They do not show that these groups *feel* the intensity of pain differentially. In fact, the most recent evaluation of all research conducted into the cultural differences in response to pain concluded:

> There is no evidence suggesting that the neuro-physiologic detection of pain (i.e., pain threshold) varies across cultural boundaries.

It is reasonable for us to conclude that pain is not differentially distributed in society according to social class or race, as is money. Certainly, if there were any differences, these would be *comparatively* much less than are those of economics. Pain is a great leveller.

In sum, people's physiological reaction to painful stimuli is pretty much the same. The way they deal

with the pain varies according to the way they have been brought up. In other words, all people feel pain as pain. The ways they react to this pain may vary.

DIFFERENCES IN REACTION TO CHRONIC PAIN

Are there differential responses to the effects of prison (that is, chronic pain and pain tolerance) according to social and ethnic background? There are drastic variations. Not only do people (including inmates) perceive time differently (and it is time that is the main element of pain in prison), but inmates also experience prison life in widely differing ways. . . .

THE FAIRNESS OF ACUTE PAIN

One can immediately see the inherently attractive features of acute pain such as electric shock, as opposed to prison. It ensures that all persons receive the same amount of punishment. All people, rich or poor, black or white, will suffer the same amount of pain. This surely fulfills the requirements of equity and fairness. People will truly receive the same amount of punishment for the same crimes. No longer will it be possible to claim that the punishment favors the rich or poor, since we know that we have, by the scientific selection of an intensely painful stimulus, ensured that each individual will experience the same amount of pain. And for those purists who would insist that no matter at what level of intensity of shock, each one will feel it differently, one may reply that even if this is so, it is demonstrably clear that in comparison with the punishment of prison, the application of physically acute pain to the body is far more equitable, and far less susceptible to variations in effects. It achieves its object, then stops.

WOULD MINORITIES SUFFER MORE THAN THEY DO NOW?

The number of people in prison has increased over the last twenty years at an astronomical rate, from approximately 177,113 in 1971 to 1.5 million in 1994.

The number of minorities in state prisons has always been disproportionately high: 47.8% in 1979, during a period when there was supposedly more sensitivity to the plight of minorities in the United States. The U.S. Bureau of Justice Statistics reports that in 1993 minorities made up 61.7% of the prison population (44.1% blacks and 17.6% hispanics). When we consider that blacks comprise only about 12% of the total United States population, we see that the proportion of blacks in prison is tremendous. And in some regions of the United States the proportions are even higher. The chances are that every black in the country has at least one relative in prison and probably more. . . .

Clearly, if corporal punishment can become a viable alternative to prison, then blacks stand to gain more than any other group. A reviewer of the first edition of this book castigated me for making this statement. But I stand by it. The biggest challenge to minorities has always been to convince the American public that they are being treated badly. For the entire period I have been involved in criminal justice (longer than I would care to admit) blacks have complained of discrimination and abuse at the hands of the police. Not until the highly public Rodney King beating did this issue become a public (and political) cause. But it is impossible to publicize the tremendous suffering caused by imprisonment because it is by design administered *in secret.* If blacks were punished in public with corporal punishment to the differential extent that they are currently punished with prison, there would be a public outcry. It would be *too much.* It would force us to be accountable for our excesses in punishment. Right now, we are not accountable for the excesses of prison, because it is a secret and silent punishment. The differential rate of black imprisonment is a *silent statistic.*

WHAT ABOUT WOMEN AND CHILDREN?

It is a popular belief that women are able to withstand more pain than men, although there is no research data to support this claim.

It is also argued on occasion that the young could withstand severe acute pain more easily than the old, but once again there is no research data to support this claim.

But there is evidence to show that women and children suffer more than do adult men from the punishment of prison. Studies have shown that women suffer the separation from their families much more than do men, and it is a well established fact that the young who are sent to prison are those who are preyed on by rapists, and if they are not raped, they are turned into hardened criminals by older inmates.

Therefore, it would seem to be much more preferable to administer acute pain in the same quantities for women and children as for men. In this way we have truly fair punishments, and all, regardless of race, age, or sex, receive the same penalty. Isn't this politically correct? In addition, the law could be considerably simplified, since in many cases, separate laws have had to be made for children so that they could be given more lenient (and recently more severe) punishments. . . .

I strongly oppose the use of corporal punishment on children within the family, because they have no recourse. It is punishment administered in secret (the sanctity of the home) and thus full of the same dangers of excess as found in the use of corporal punishment in prisons. There is no ready accountability, no "due process" one might say, within the traditional family household. Corporal punishment is a very severe punishment. It should only be used against those who have broken the law. There is no justification for its use otherwise.

IN SUM . . .

It is clear that, where appropriate to the crimes, the use of an acute corporal punishment instead of the vague use of prison as punishment is preferable, since the application of acute pain can be scientifically controlled both in terms of duration and intensity. The way it is physically felt does not vary, although reactions to it may: but even these reactions vary far less than reactions to prison.

Prisons conditions vary so much, that there is no reliable way to control their quality or intensity. Convicts are well aware of the ways in which prisons differ from each other, as is well documented in any prison diaries and books about prison. And these variations occur within prisons that are supposed to be of one type, such as "maximum security."

This chapter suggests another defect in the use of prison as punishment. It is difficult to vary the intensity and duration of prison in a clear-cut way. The most common way to vary intensity of prison is to vary the length of prison term. But this mixes up duration with intensity. Another way to vary the intensity of prison is to introduce various types or degrees of "security"—such as maximum security down to minimum security prisons. However, such variations in intensity are not specific enough for our purposes, for we wish to control the administration of pain as carefully and fully as possible.

Varying the intensity of prisons also affects the credibility of prisons as a punishment since minimum security prisons are easily portrayed as "resorts for white collar criminals." "Perks" such as gym equipment, TV, or even libraries, convey the (mistaken) idea that prison is not that bad a place. Varying the intensity of prison in this way undermines its credibility as a punishment, and paradoxically feeds the public's demand for more and more punishment, which under the present system can only be prison.

We have seen that we can minutely control both the intensity and duration of electric shock. If we are to control the intensity of prison, we must look closely at the types of pain that occur in prison—diet, hard labor, isolation cells—and consider systematically grading these so that the intensity of prison may be adjusted to the punishment deserved by the crime.

REVIEW AND DISCUSSION QUESTIONS

1. What's the difference between pain threshold and tolerance of pain? Is it true that people may react to pain differently, but that their physiological experience of it is pretty much the same? Does Graeme Newman respond adequately to the problem of making the pain of electric shock the same for everyone? Is inequality of pain also a problem for the use of prison as punishment?
2. What are Newman's main reasons for advocating corporal punishment—in particular, the use of electric shock—over prison? Assess each of the reasons he gives. How persuasive do you find them?

3. Would corporal punishment be more effective in controlling crime than prison is, or less effective? Would it be more or less fair than prison? In general, what do you see as the advantages and disadvantages of Newman's proposal? Assess the use of corporal punishment versus imprisonment from the point of view of each of the theories of punishment presented in Section 12 of this book.
4. Why does Newman believe that blacks would gain more than any other group from the use of corporal punishment? Explain why you agree or disagree.
5. Should corporeal punishment be inflicted on children either by the law (a practice Newman supports) or inside the family (opposed by Newman)?
6. Assess the contention that the use of electric shock would be a "cruel and unusual" punishment and thus in violation of the Eighth Amendment.
7. Assuming Newman's proposal is basically sound, are there crimes for which corporal punishment would be inappropriate?

"Three Strikes"

Rummel v. Estelle

In 1964, William James Rummel was convicted of fraudulent use of a credit card, in 1969 of forging a check, and again in 1973 of obtaining money under false pretenses. The sums involved were $80, $28.36, and $120.75, respectively. Like some other states, Texas has a recidivist or "three strikes and you're out" statute, which directs that a criminal convicted of a third felony be sentenced to life imprisonment. In this case, the Supreme Court examines the question of whether putting Rummel in prison for life violates the Eighth Amendment's ban on "cruel and unusual" punishment.

Mr. Justice Rehnquist delivered the opinion of the Court: Petitioner William James Rummel is presently serving a life sentence imposed by the State of Texas in 1973 under its "recidivist statute," . . . which provided that "[w]hoever shall have been three times convicted of a felony less than capital shall on such third conviction be imprisoned for life in the penitentiary." On January 19, 1976, Rummel sought a writ of habeas corpus in the United States District Court for the Western District of Texas, arguing that life imprisonment was "grossly disproportionate" to the three felonies that formed the predicate for his sentence and that therefore the sentence violated the ban on cruel and unusual punishments of the Eighth and Fourteenth Amendments. The District Court and the United States Court of Appeals for the Fifth Circuit rejected Rummel's claim, finding no unconstitutional disproportionality. We granted certiorari . . . and now affirm.

In 1964 the State of Texas charged Rummel with fraudulent use of a credit card to obtain $80 worth of goods or services. Because the amount in question was greater than $50, the charged offense was a felony punishable by a minimum of 2 years and a maximum of 10 years in the Texas Department of Corrections. Rummel eventually pleaded guilty to the charge and was sentenced to three years' confinement in a state penitentiary.

In 1969 the State of Texas charged Rummel with passing a forged check in the amount of $28.36, a crime punishable by imprisonment in a penitentiary for not less than two nor more than five years. Rummel pleaded guilty to this offense and was sentenced to four years' imprisonment.

445 U.S. 263 (1980). Some notes and citations omitted.

In 1973 Rummel was charged with obtaining $120.75 by false pretenses. Because the amount obtained was greater than $50, the charged offense was designated "felony theft," which, by itself, was punishable by confinement in a penitentiary for not less than 2 nor more than 10 years. The prosecution chose, however, to proceed against Rummel under Texas'[s] recidivist statute, and cited in the indictment his 1964 and 1969 convictions as requiring imposition of a life sentence if Rummel were convicted of the charged offense. A jury convicted Rummel of felony theft and also found as true the allegation that he had been convicted of two prior felonies. As a result, on April 26, 1973, the trial court imposed upon Rummel the life sentence mandated by Art. 63 [of the Texas Penal Code]. . . .

This Court has on occasion stated that the Eighth Amendment prohibits imposition of a sentence that is grossly disproportionate to the severity of the crime. In recent years this proposition has appeared most frequently in opinions dealing with the death penalty. . . .

Because a sentence of death differs in kind from any sentence of imprisonment, no matter how long, our decisions applying the prohibition of cruel and unusual punishments to capital cases are of limited assistance in deciding the constitutionality of the punishment meted out to Rummel.

Outside the context of capital punishment, successful challenges to the proportionality of particular sentences have been exceedingly rare. In *Weems v. United States* [217 U.S. 349 (1910)], a case coming to this Court from the Supreme Court of the Philippine Islands, petitioner successfully attacked the imposition of a punishment known as "cadena temporal" for the crime of falsifying a public record. . . . The mandatory "remedy" for this offense was *cadena temporal,* a punishment described graphically by the Court:

> Its minimum degree is confinement in a penal institution for twelve years and one day, a chain at the ankle and wrist of the offender, hard and painful labor, no assistance from friend or relative, no marital authority or parental rights or rights of property, no participation even in the family council. These parts of his penalty endure for the term of imprisonment. From other parts there is no intermission. His prison bars and chains are removed, it is true, after twelve years, but he goes from them to a perpetual limitation of his liberty. He is forever kept under the shadow of his crime, forever kept within voice and view of the criminal magistrate, not being able to change his domicile without giving notice to the "authority immediately in charge of his surveillance," and without permission in writing. . . .

Although Rummel argues that the length of Weems's imprisonment was, by itself, a basis for the Court's decision, the Court's opinion does not support such a simple conclusion. The opinion consistently referred jointly to the length of imprisonment and its "accessories" or "accompaniments." . . .

In an attempt to provide us with objective criteria against which we might measure the proportionality of his life sentence, Rummel points to certain characteristics of his offenses that allegedly render them "petty." He cites, for example, the absence of violence in his crimes. But the presence or absence of violence does not always affect the strength of society's interest in deterring a particular crime or in punishing a particular criminal. A high official in a large corporation can commit undeniably serious crimes in the area of antitrust, bribery, or clean air or water standards without coming close to engaging in any "violent" or short-term "life-threatening" behavior. Additionally, Rummel cites the "small" amount of money taken in each of his crimes. But to recognize that the State of Texas could have imprisoned Rummel for life if he had stolen $5,000, $50,000, or $500,000, rather than the $120.75 that a jury convicted him of stealing, is virtually to concede that the lines to be drawn are indeed "subjective," and therefore properly within the province of legislatures, not courts. Moreover, if Rummel had attempted to defraud his victim of $50,000, but had failed, no money whatsoever would have changed hands; yet Rummel would be no less blameworthy, only less skillful, than if he had succeeded.

In this case, however, we need not decide whether Texas could impose a life sentence upon Rummel merely for obtaining $120.75 by false pretenses. Had Rummel only committed that crime, under the law enacted by the Texas Legislature he could have been imprisoned for no more than 10 years. In fact, at the time that he obtained the $120.75 by false pretenses, he already had committed and had been imprisoned for two other felonies, crimes that Texas and other

States felt were serious enough to warrant significant terms of imprisonment even in the absence of prior offenses. Thus the interest of the State of Texas here is not simply that of making criminal the unlawful acquisition of another person's property; it is in addition the interest, expressed in all recidivist statutes, in dealing in a harsher manner with those who by repeated criminal acts have shown that they are simply incapable of conforming to the norms of society as established by its criminal law. By conceding the validity of recidivist statutes generally, Rummel himself concedes that the State of Texas, or any other State, has a valid interest in so dealing with that class of persons.

. . . Thus, under Art. 63, a three-time felon receives a mandatory life sentence, with possibility of parole, only if commission and conviction of each succeeding felony followed conviction for the preceding one, and only if each prior conviction was followed by actual imprisonment. Given this necessary sequence, a recidivist must twice demonstrate that conviction and actual imprisonment do not deter him from returning to crime once he is released. One in Rummel's position has been both graphically informed of the consequences of lawlessness and given an opportunity to reform, all to no avail. Article 63 thus is nothing more than a societal decision that when such a person commits yet another felony, he should be subjected to the admittedly serious penalty of incarceration for life, subject only to the State's judgment as to whether to grant him parole.[1]

Mr. Justice Powell, with whom Mr. Justice Brennan, Mr. Justice Marshall, and Mr. Justice Stevens join, dissenting: The scope of the Cruel and Unusual Punishments Clause extends not only to barbarous methods of punishment, but also to punishments that are grossly disproportionate. Disproportionality analysis measures the relationship between the nature and number of offenses committed and the severity of the punishment inflicted upon the offender. The inquiry focuses on whether a person deserves such punishment, not simply on whether punishment would serve a utilitarian goal. A statute that levied a mandatory life sentence for overtime parking might well deter vehicular lawlessness, but it would offend our felt sense of justice. The Court concedes today that the principle of disproportionality plays a role in the review of sentences imposing the death penalty, but suggests that the principle may be less applicable when a noncapital sentence is challenged. Such a limitation finds no support in the history of Eighth Amendment jurisprudence.

The principle of disproportionality is rooted deeply in English constitutional law. The Magna Carta of 1215 insured that "[a] free man shall not be [fined] for a trivial offence, except in accordance with the degree of the offence; and for a serious offence he shall be [fined] according to its gravity." By 1400, the English common law had embraced the principle, not always followed in practice, that punishment should not be excessive either in severity or length. . . .

In sum, a few basic principles emerge from the history of the Eighth Amendment. Both barbarous forms of punishment and grossly excessive punishments are cruel and unusual. A sentence may be excessive if it serves no acceptable social purpose, or is grossly disproportionate to the seriousness of the crime. The principle of disproportionality has been acknowledged to apply to both capital and noncapital sentences. . . .

Examination of the objective factors traditionally employed by the Court to assess the proportionality of a sentence demonstrates that petitioner suffers a cruel and unusual punishment. Petitioner has been sentenced to the penultimate criminal penalty because he committed three offenses defrauding others of about $230. The nature of the crimes does not suggest that petitioner ever engaged in conduct that threatened another's person, involved a trespass, or endangered in any way the peace of society. A comparison of the sentence petitioner received with the sentences provided by habitual offender statutes of other American jurisdictions demonstrates that only two other States authorize the same punishment. A comparison of petitioner to other criminals sentenced in Texas shows that he has been punished for three property-related offenses with a harsher sentence than that given first-time offenders or two-time offenders convicted of far more serious offenses. The Texas system assumes that all three-time offenders deserve the same punishment whether they commit three murders or cash three fraudulent checks.

The petitioner has committed criminal acts for which he may be punished. He has been given a sentence that is not inherently barbarous. But the

relationship between the criminal acts and the sentence is grossly disproportionate. For having defrauded others of about $230, the State of Texas has deprived petitioner of his freedom for the rest of his life. The State has not attempted to justify the sentence as necessary either to deter other persons or to isolate a potentially violent individual. Nor has petitioner's status as a habitual offender been shown to justify a mandatory life sentence. My view, informed by examination of the "objective indicia that reflect the public attitude toward a given sanction," is that this punishment violates the principle of proportionality contained within the Cruel and Unusual Punishments Clause.

NOTE

1. Thus, it is not true that, as the dissent claims, the Texas scheme subjects a person to life imprisonment "merely because he is a three-time felon." On the contrary, Art. 63 mandates such a sentence only after shorter terms of actual imprisonment have proved ineffective.

REVIEW AND DISCUSSION QUESTIONS

1. On what basis does the Court reject the claim that Rummel's sentence was grossly disproportionate to the severity of his crime? What are the dissenters' arguments to the contrary? With whom do you agree and why? Do you think Rummel's sentence was unconstitutional?
2. In your view, were Rummel's offenses "petty"? Do you think the Texas legislature had criminals like Rummel in mind when it drafted its recidivist statute?
3. Constitutionality aside, are such statutes wise social policy? Would a utilitarian reject them, defend them, or modify them in some way? Because one clearly does not deserve life imprisonment merely for obtaining $120 on false pretenses, is there any way a retributivist could justify sentencing Rummel to life in prison?
4. Could recidivist statutes be defended, not as punishment, but rather as a way of incapacitating an individual who has shown himself or herself to be dangerous to society?
5. Are recidivist statutes fair? If two people commit the same crime (or even commit a crime together), would one of them deserve a stiffer sentence simply because he or she is a previous offender?

Racial Bias in Sentencing

McCleskey v. Kemp

Some studies contend that blacks are more likely to receive the death penalty than whites who commit the same crime. Other studies challenge this conclusion. But there is no dispute about figures involving the impact of the race of the victim: People who kill whites are more likely to be executed than people who kill blacks. The Supreme Court considered the legal implications of this fact in the *McCleskey* case. A jury of eleven whites and one black convicted Warren McCleskey, who is black, of murdering a white man. McCleskey appealed his conviction, relying on a detailed study of Georgia's sentencing patterns conducted by David Baldus and others. The Baldus study indicated that anybody, black or white, who kills a white is 4.3 times more likely to receive the death penalty than someone who kills a black. McCleskey's lawyers used this and other figures to argue that Georgia's criminal justice system is racially biased and that, therefore, it denied McCleskey equal protection of the law.

---

107 S. Ct. 1756 (1987). Footnotes omitted.

JUSTICE POWELL delivered the opinion of the Court: Out analysis begins with the basic principle that a defendant who alleges an equal protection violation has the burden of proving "the existence of purposeful discrimination." . . . McCleskey must prove that the decisionmakers in *his* case acted with discriminatory purpose. He offers no evidence specific to his own case that would support an inference that racial considerations played a part in his sentence. Instead, he relies solely on the Baldus study. . . .

Each jury is unique in its composition, and the Constitution requires that its decision rest on consideration of innumerable factors that vary according to the characteristics of the individual defendant and the facts of the particular capital offense. . . .

Implementation of laws necessarily requires discretionary judgments. Because discretion is essential to the criminal justice process, we would demand exceptionally clear proof before we would infer that the discretion has been abused. . . .

McCleskey also suggests that the Baldus study proves that the State as a whole has acted with a discriminatory purpose. . . . [In *Furman*] this Court found that the Georgia capital sentencing system could operate in a fair and neutral manner. There was no evidence then, and there is none now, that the Georgia Legislature enacted the capital punishment statute to further a racially discriminatory purpose. . . .

[McCleskey] contends that the Georgia capital punishment system is arbitrary and capricious in a *application,* and therefore this sentence is excessive, because racial considerations may influence capital sentencing decisions in Georgia. . . .

The capital sentencing decision requires the individual jurors to focus their collective judgment on the unique characteristics of a particular criminal defendant. It is not surprising that such collective judgments often are difficult to explain. But the inherent lack of predictability of jury decisions does not justify their condemnation. On the contrary, it is the jury's function to make the difficult and uniquely human judgments that defy codification and that "buil[d] discretion, equity, and flexibility into a legal system." . . .

Even Professor Baldus does not contend that his statistics *prove* that race enters into any capital sentencing decisions or that race was a factor in McCleskey's particular case. Statistics at most may show only a likelihood that a particular factor entered into some decisions. There is, of course, some risk of racial prejudice influencing a jury's decision in a criminal case. . . . The question is at what point that risk becomes constitutionally unacceptable. At most, the Baldus study indicates a discrepancy that appears to correlate with race. Apparent disparities in sentencing are an inevitable part of our criminal justice system. . . . As this Court has recognized, any mode for determining guilt or punishment "has its weaknesses and the potential for misuse." . . .

Where the discretion that is fundamental to our criminal process is involved, we decline to assume that what is unexplained is invidious. In light of the safeguards designed to minimize racial bias in the process, the fundamental value of jury trial in our criminal justice system, and the benefits that discretion provides to criminal defendants, we hold that the Baldus study does not demonstrate a constitutionally significant risk of racial bias affecting the Georgia capital-sentencing process.

JUSTICE BRENNAN, dissenting: At some point in this case, Warren McCleskey doubtless asked his lawyer whether a jury was likely to sentence him to die. A candid reply to this question would have been disturbing. First, counsel would have to tell McCleskey that few of the details of the crime or of McCleskey's past criminal conduct were more important than the fact that his victim was white. Furthermore, counsel would feel bound to tell McCleskey that defendants charged with killing white victims in Georgia are 4.3 times as likely to be sentenced to death as defendants charged with killing blacks. In addition, frankness would compel the disclosure that it was more likely than not that the race of McCleskey's victim would determine whether he received a death sentence: 6 of every 11 defendants convicted of killing a white person would not have received the death penalty if their victims had been black, while, among defendants with aggravating and mitigating factors comparable to McCleskey, 20 of every 34 would not have been sentenced to die if their victims had been black. Finally, the assessment would not be complete without the information that cases involving black defendants and white victims are more likely to result in a death sentence than cases featuring any other racial combination of

defendant and victim. The story could be told in a variety of ways, but McCleskey could not fail to grasp its essential narrative line: there was a significant chance that race would play a prominent role in determining if he lived or died.

The Court today holds that Warren McCleskey's sentence was constitutionally imposed. It finds no fault in a system in which lawyers must tell their clients that race casts a large shadow on the capital sentencing process. The Court arrives at this conclusion by stating that the Baldus Study cannot "*prove* that race enters into any capital sentencing decisions or that race was a factor in McCleskey's particular case." According to Professor Baldus, . . . we can identify only "a likelihood that a particular factor entered into some decisions," and "a discrepancy that appears to correlate with race." This "likelihood" and "discrepancy," holds the Court, is insufficient to establish a constitutional violation. . . .

The Baldus study indicates that, after taking into account some 230 nonracial factors that might legitimately influence a sentencer, the jury *more likely than not* would have spared McCleskey's life had his victim been black. . . . Data unadjusted for the mitigating or aggravating effect of other factors show an even more pronounced disparity by race. The capital sentencing rate for all white-victim cases was almost *11 times* greater than the rate for black-victim cases. Furthermore, blacks who kill whites are sentenced to death at nearly *22 times* the rate of blacks who kill blacks, and more than *7 times* the rate of whites who kill blacks. In addition, prosecutors seek the death penalty for 70% of black defendants with white victims, but for only 15% of black defendants with black victims, and only 19% of white defendants with black victims. . . .

Georgia's legacy of a race-conscious criminal justice system, as well as this Court's own recognition of the persistent danger that racial attitudes may affect criminal proceedings, indicate that McCleskey's claim is not a fanciful product of mere statistical artifice. . . .

For many years, Georgia operated openly and formally precisely the type of dual system the evidence shows is still effectively in place. The criminal law expressly differentiated between crimes committed by and against blacks and whites, distinctions whose lineage traced back to the time of slavery. . . .

This Court has invalidated portions of the Georgia capital sentencing system 3 times over the past 15 years. The specter of race discrimination was acknowledged by the Court in striking down the Georgia death-penalty statute in *Furman*. . . .

Formal dual criminal laws may no longer be in effect, and intentional discrimination may no longer be prominent. Nonetheless, "subtle, less consciously held racial attitudes" continue to be of concern, and the Georgia system gives such attitudes considerable room to operate. . . . The determination of the significance of his evidence is at its core an exercise in human moral judgment, not a mechanical statistical analysis. . . . It is true that every nuance of decision cannot be statistically captured, nor can any individual judgment be plumbed with absolute certainty. Yet the fact that we must always act without the illumination of complete knowledge cannot induce paralysis when we confront what is literally an issue of life and death.

REVIEW AND DISCUSSION QUESTIONS

1. Why does the Court reject the statistical argument of McCleskey's lawyers? What would McCleskey have to establish to show that he had been denied equal protection? For what reasons does Brennan dissent?
2. McCleskey's claim was not that there was prejudice against him. Rather, racial bias is alleged to operate at one remove, that is, he probably would not have been sentenced to death if his *victim* had been black. Has McCleskey therefore been treated unfairly? Under what circumstances, if any, can an individual sentence be just if the criminal justice system as a whole is infected by racism?
3. Defenders of capital punishment argue that those who murder blacks are getting off too easily while those who murder whites are simply getting what they (and all murderers) deserve. The real racial discrimination, they contend, is not against criminals like McCleskey but against their black victims, whose murderers get off too easily. Assess this argument.
4. Can capital punishment be morally or legally justified if it is administered in an arbitrary or racially biased way? Can any punishment?

CHAPTER 15

Criminal Responsibility

Even if we assume that punishment is morally justifiable, that the rights of defendants are respected, and that a legal system's sentencing policies are effective and fair, many intriguing and philosophically important issues are raised by the law's effort to specify the exact extent of one's liability to criminal prosecution. In particular, when should conduct that is normally illegal or wrongful be excused or its punishment mitigated?

The Queen v. Dudley and Stephens poses this issue in a very dramatic context, namely, when killing another person appears to be the only way to save one's own life. In "The Principles of Criminal Law," Richard B. Brandt explains the legal elements of a crime and distinguishes different kinds of legal justifications and excuses, while H. L. A. Hart probes the important role of intention in criminal law. Leo Katz then examines criminal liability for crimes that were impossible for the accused person to have committed. *People v. Koerber* focuses on intoxication and criminal intent, while *State v. Leidholm* takes up self-defense and the battered woman syndrome. The case of *Regina v. Morgan* and the essay by Susan Estrich deal with the issue of criminal responsibility in cases of rape. The next two essays discuss the insanity defense. R. J. Gerber proposes restructuring the legal treatment of insanity, and Joel Feinberg explores the moral significance of mental illness with respect to blame and punishment. The section concludes with *Atkins v. Virginia*, in which the Supreme Court rules on the constitutionality of executing mentally retarded criminals.

Survival on a Lifeboat

The Queen v. Dudley and Stephens

Life-and-death decisions in extreme circumstances make for hard cases, morally and legally. This famous English case concerns four sailors adrift in a lifeboat. After twenty days with little food or water, two of them—Dudley and Stephens—killed the youngest member of their party, a boy seventeen or eighteen years old, and ate him. Four days later, the sailors were rescued. It is unlikely that they would have survived otherwise, and it is probable that the boy would have died before them.

Indictment for the murder of Richard Parker on the high seas within the jurisdiction of the Admiralty: At the trial before Huddleston, B., at the Devon and Cornwall Winter Assizes, November 7, 1884, the jury, at the suggestion of the learned judge, found the facts of the case in a special verdict which stated "that on July 5, 1884, the prisoners, Thomas Dudley and Edward Stephens, with one Brooks, all able bodied English seamen, and the deceased also an English boy, between seventeen and eighteen years of age, the crew of an English yacht, a registered English vessel, were cast away in a storm on the high seas 1600 miles from the Cape of Good Hope, and were compelled to put into an open boat belonging to the said yacht. That in this boat they had no supply of water and no supply of food, except two 1 lb. tins of turnips, and for three days they had nothing else to subsist upon. That on the fourth day they caught a small turtle, upon which they subsisted for a few days, and this was the only food they had up to the twentieth day when the act now in question was committed. That on the twelfth day the remains of the turtle were entirely consumed, and for the next eight days they had nothing to eat. That they had no fresh water, except such rain as they from time to time caught in their oilskin capes. That the boat was drifting on the ocean, and was probably more than 1000 miles away from land. That on the eighteenth day, when they had been seven days without food and five without water, the prisoners spoke to Brooks as to what should be done if no succour came, and suggested that some one should be sacrificed to save the rest, but Brooks dissented, and the boy, to whom they were understood to refer, was not consulted. That on the 24th of July, the day before the act now in question, the prisoner Dudley proposed to Stephens and Brooks that lots should be cast who should be put to death to save the rest, but Brooks refused to consent, and it was not put to the boy, and in point of fact there was no drawing of lots. That on that day the prisoners spoke of their having families, and suggested it would be better to kill the boy that their lives should be saved, and Dudley proposed that if there was no vessel in sight by the morrow morning the boy should be killed. That next day, the 25th of July, no vessel appearing, Dudley told Brooks that he had better go and have a sleep, and made signs to Stephens and Brooks that the boy had better be killed. The prisoner Stephens agreed to the act, but Brooks dissented from it. That the boy was then lying at the bottom of the boat quite helpless, and extremely weakened by famine and by drinking sea water, and unable to make any resistance, nor did he ever assent to his being killed. The prisoner Dudley offered a prayer asking forgiveness for them all if either of them should be tempted to commit a rash act, and that their souls might be saved. That Dudley, with the assent of Stephens, went to the boy, and telling him that his time was come, put a knife into his throat and killed him then and there; that the three men fed upon the body and blood of the boy for four days; that on the fourth day after the act had been committed the boat was picked up by a passing vessel, and the prisoners were rescued, still alive, but in the lowest state of prostration. That they were carried to the port of Falmouth, and committed for trial at Exeter. That if the men had not fed upon the body of the boy they would probably not have survived to be so picked up and rescued, but would within the four days have died of famine. That the boy, being in a much weaker condition, was likely to have died before them. That at the time of the act in question there was no sail in sight, nor any reasonable prospect of relief. That under these circumstances there appeared to the prisoners every probability that unless they then fed or very soon fed upon the boy or one of themselves they would die of starvation. That there was no appreciable chance of saving life except by killing some one for the others to eat. That assuming any necessity to kill anybody, there was no greater necessity for killing the boy than any of the other three men. But whether upon the whole matter by the jurors found the killing of Richard Parker by Dudley and Stephens be felony and murder the jurors are ignorant, and pray the advice of the Court thereupon, and if upon the whole matter the Court shall be of opinion that the killing of Richard Parker be felony and murder, then the jurors say that Dudley and Stephens were each guilty of felony and murder as alleged in the indictment."

The learned judge then adjourned the assizes until the 25th of November at the Royal Courts of Justice. On the application of the Crown they were again adjourned to the 4th of December, and the case ordered to be argued before a Court consisting of five judges.

Dec. 4. Sir H. James, A.G. appeared [as prosecutor] for the Crown: . . . With regard to the substantial question in the case—whether the prisoners in killing Parker were guilty of murder—the law is that where a private person acting upon his own judgment takes the life of a fellow creature, his act can only be justified on the ground of self-defence—self-defence against the acts of the person whose life is taken. This principle has been extended to include the case of a man killing another to prevent him from committing some great crime upon a third person. But the principle has no application to this case, for the prisoners were not protecting themselves against any act of Parker. If he had had food in his possession and they had taken it from him, they would have been guilty of theft; and if they killed him to obtain this food, they would have been guilty of murder. . . .

A. Collins, Q.C., for the prisoners: The facts found on the special verdict shew that the prisoners were not guilty of murder, at the time when they killed Parker, but killed him under the pressure of necessity. Necessity will excuse an act which would otherwise be a crime. Stephen, *Digest of Criminal Law,* art. 32, Necessity. The law as to compulsion by necessity is further explained in Stephen's *History of the Criminal Law,* vol. ii., p. 108, and an opinion is expressed that in the case often put by casuists, of two drowning men on a plank large enough to support one only, and one thrusting the other off, the survivor could not be subjected to legal punishment. In the American case of *The United States v. Holmes,* the proposition that a passenger on board a vessel may be thrown overboard to save the others is sanctioned. The law as to inevitable necessity is fully considered in *Russell on Crimes,* vol. i. p. 847. . . . Lord Bacon . . . gives the instance of two shipwrecked persons clinging to the same plank and one of them thrusting the other from it, finding that it will not support both, and says that this homicide is excusable through unavoidable necessity and upon the great universal principle of self-preservation, which prompts every man to save his own life in preference to that of another, where one of them must inevitably perish. It is true that Hale's *Pleas of the Crown,* p. 54, states distinctly that hunger is no excuse for theft, but that is on the ground that there can be no such extreme necessity in this country. In the present case the prisoners were in circumstances where no assistance could be given. The essence of the crime of murder is intention, and here the intention of the prisoners was only to preserve their lives. . . .

Dec. 9. The judgment of the Court was delivered by Lord Coleridge, C.J.: The two prisoners, Thomas Dudley and Edwin Stephens, were indicted for the murder of Richard Parker on the high seas on the 25th of July in the present year. They were tried before my Brother Huddleston at Exeter on the 6th of November, and, under the direction of my learned Brother, the jury returned a special verdict, the legal effect of which has been argued before us, and on which we are now to pronounce judgment.

The special verdict as, after certain objections by Mr. Collins to which the Attorney General yielded, it is finally settled before us is as follows. [His Lordship read the special verdict as above set out.] From these facts, stated with the cold precision of a special verdict, it appears sufficiently that the prisoners were subject to terrible temptation, to sufferings which might break down the bodily power of the strongest man, and try the conscience of the best. Other details yet more harrowing, facts still more loathsome and appalling, were presented to the jury, and are to be found recorded in my learned Brother's notes. But nevertheless this is clear, that the prisoners put to death a weak and unoffending boy upon the chance of preserving their own lives by feeding upon his flesh and blood after he was killed, and with the certainty of depriving *him* of any possible chance of survival. The verdict finds in terms that "if the men had not fed upon the body of the boy they would *probably* not have survived," and that "the boy being in a much weaker condition was *likely* to have died before them." They might possibly have been picked up next day by a passing ship; they might possibly not have been picked up at all; in either case it is obvious that the killing of the boy would have been an unnecessary and profitless act. It is found by the verdict that the boy was incapable of resistance, and, in fact, made none; and it is not even suggested that his death was due to any violence on his part attempted against, or even so much as feared by, those who killed him. Under

these circumstances the jury say that they are ignorant whether those who killed him were guilty of murder, and have referred it to this Court to determine what is the legal consequence which follows from the facts which they have found. . . .

There remains to be considered the real question in the case—whether killing under the circumstances set forth in the verdict be or be not murder. The contention that it could be anything else was, to the minds of us all, both new and strange, and we stopped the Attorney General in his negative argument in order that we might hear what could be said in support of a proposition which appeared to us to be at once dangerous, immoral, and opposed to all legal principle and analogy. All, no doubt, that can be said has been urged before us, and we are now to consider and determine what it amounts to. First it is said that it follows from various definitions of murder in books of authority, which definitions imply, if they do not state, the doctrine, that in order to save your own life you may lawfully take away the life of another, when that other is neither attempting nor threatening yours, nor is guilty of any illegal act whatever towards you or any one else. But if these definitions be looked at they will not be found to sustain this contention. . . .

Is there, then, any authority for the proposition which has been presented to us? Decided cases there are none. . . . The American case cited by my Brother Stephen in his *Digest,* from *Wharton on Homicide,* in which it was decided, correctly indeed, that sailors had no right to throw passengers overboard to save themselves, but on the somewhat strange ground that the proper mode of determining who was to be sacrificed was to vote upon the subject by ballot, can hardly, as my Brother Stephen says, be an authority satisfactory to [this court]. . . .

The one real authority of former time is Lord Bacon, who . . . lays down the law as follows:—"Necessity carrieth a privilege in itself. Necessity is of three sorts—necessity of conservation of life, necessity of obedience, and necessity of the act of God or of a stranger. First of conservation of life; if a man steal viands to satisfy his present hunger, this is no felony nor larceny. So if divers be in danger of drowning by the casting away of some boat or barge, and one of them get to some plank, or on the boat's side to keep himself above water, and another to save his life thrust him from it, whereby he is drowned, this is neither se defendendo nor by misadventure, but justifiable." On this it is to be observed that Lord Bacon's proposition that stealing to satisfy hunger is no larceny is hardly supported by Staundforde, whom he cites for it, and is expressly contradicted by Lord Hale. . . . And for the proposition as to the plank or boat, it is said to be derived from the canonists. At any rate he cites no authority for it, and it must stand upon his own. Lord Bacon was great even as a lawyer; but it is permissible to much smaller men, relying upon principle and on the authority of others, the equals and even the superiors of Lord Bacon as lawyers, to question the soundness of his dictum. There are many conceivable states of things in which it might possibly be true, but if Lord Bacon meant to lay down the broad proposition that a man may save his life by killing, if necessary, an innocent and unoffending neighbour, it certainly is not law at the present day.

There remains the authority of my Brother Stephen, who, both in his *Digest* and in his *History of the Criminal Law,* uses language perhaps wide enough to cover this case. The language is somewhat vague in both places, but it does not in either place cover this case of necessity, and we have the best authority for saying that it was not meant to cover it.

. . . We are dealing with a case of private homicide, not one imposed upon men in the service of their Sovereign and in the defence of their country. Now it is admitted that the deliberate killing of this unoffending and unresisting boy was clearly murder, unless the killing can be justified by some well-recognised excuse admitted by the law. It is further admitted that there was in this case no such excuse, unless the killing was justified by what has been called "necessity." But the temptation to the act which existed here was not what the law has ever called necessity. Nor is this to be regretted. Though law and morality are not the same, and many things may be immoral which are not necessarily illegal, yet the absolute divorce of law from morality would be of fatal consequence; and such divorce would follow if the temptation to murder in this case were to be held by law an absolute defence of it. It is not so. To preserve one's life is generally speaking a duty, but it may be the plainest and the highest duty to sacrifice it. War is full of instances in which it is a man's duty not to live, but to die.

The duty, in case of shipwreck, of a captain to his crew, of the crew to the passengers, of soldiers to women and children, as in the noble case of the *Birkenhead;* these duties impose on men the moral necessity, not of the preservation, but of the sacrifice of their lives for others, from which in no country, least of all, it is to be hoped, in England, will men ever shrink, as indeed, they have not shrunk. It is not correct, therefore, to say that there is any absolute or unqualified necessity to preserve one's life. . . . It is not needful to point out the awful danger of admitting the principle which has been contended for. Who is to be the judge of this sort of necessity? By what measure is the comparative value of lives to be measured? Is it to be strength, or intellect, or what? It is plain that the principle leaves to him who is to profit by it to determine the necessity which will justify him in deliberately taking another's life to save his own. In this case the weakest, the youngest, the most unresisting, was chosen. Was it more necessary to kill him than one of the grown men? The answer must be "No." . . . [I]t is quite plain that such a principle once admitted might be made the legal cloak for unbridled passion and atrocious crime. There is no safe path for judges to tread but to ascertain the law to the best of their ability and to declare it according to their judgment; and if in any case the law appears to be too severe on individuals, to leave it to the Sovereign to exercise that prerogative of mercy which the Constitution has intrusted to the hands fittest to dispense it.

It must not be supposed that in refusing to admit temptation to be an excuse for crime it is forgotten how terrible the temptation was; how awful the suffering; how hard in such trials to keep the judgment straight and the conduct pure. We are often compelled to set up standards we cannot reach ourselves, and to lay down rules which we could not ourselves satisfy. But a man has no right to declare temptation to be an excuse, though he might himself have yielded to it, nor allow compassion for the criminal to change or weaken in any manner the legal definition of the crime. It is therefore our duty to declare that the prisoners' act in this case was wilful murder, that the facts as stated in the verdict are no legal justification of the homicide; and to say that in our unanimous opinion the prisoners are upon this special verdict guilty of murder.[1]

The Court then proceeded to pass sentence of death upon the prisoners.[2]

NOTES

1. My brother Grove has furnished me with the following suggestion, too late to be embodied in the judgment but well worth preserving: "If the two accused men were justified in killing Parker, then if not rescued in time, two of the three survivors would be justified in killing the third, and of the two who remained the stronger would be justified in killing the weaker, so that three men might be justifiably killed to give the fourth a chance of surviving." [Note by Lord Coleridge.]
2. This sentence was afterwards commuted by the Crown to six months' imprisonment [Eds.].

REVIEW AND DISCUSSION QUESTIONS

1. Counsel for the Crown (i.e., the prosecutor) says that only self-defense can justify the taking of a life. The defense lawyer says that necessity can excuse a crime. What is the difference between justifying an action and excusing it?
2. Does "necessity" excuse, either morally or legally, an action that would otherwise be a crime, as the lawyer for Dudley and Stephens claims? If not, why not? If so, what constitutes "necessity," and when, if ever, can necessity excuse the killing of a human being? How would a utilitarian assess the defense of necessity?
3. In finding Dudley and Stephens guilty of murder, the Court says that there was "temptation" but not "necessity" and, moreover, that there is no "absolute or unqualified necessity to preserve one's life." Do you agree?

4. Assess the Court's contention that to excuse Dudley and Stephens would set a dangerous precedent.
5. Dudley and Stephens were sentenced to death for murder, but their sentence was later commuted to six months of imprisonment. This strikes many people as just. But if they were really guilty of murder, should they have been imprisoned for only six months? And if they were not guilty, why should they be imprisoned at all?
6. What verdict would you have given if you were judge? What would you have done if you were in the lifeboat?
7. Is it reasonable to apply standard moral and legal principles to a case like this, or are morality and law inapplicable in such extreme circumstances?

The Principles of Criminal Law

Richard B. Brandt

In this selection, Richard Brandt reviews the main concepts and principles of criminal law, thus providing the legal and theoretical backdrop to the issues discussed in this section. After comparing moral and legal obligation, he carefully distinguishes different kinds of legal justifications and excuses and goes on to explore the ways in which legal guilt and moral reprehensibility can diverge.

Let us review the main concepts and principles of criminal law. Unfortunately there is controversy among judges and professors of law about what these concepts and principles are, and about whether they should be changed; our brief survey must necessarily ignore such differences of opinion and other subtleties, in large part.

The central feature of the system of criminal justice in the United States is the existence of laws, statutory or common, requiring that persons be caused pain or loss by the state if a judicial process has found them guilty of a crime. It is difficult to explain in general what constitutes a crime, but we can, for a start, say that criminal action always includes overt behavior which causes or threatens a "public harm," that is, some effects considered to be harmful to the community as a whole. The law provides different penalties for different offenses, the more severe ones attaching to those crimes considered in some sense more serious. For most offenses the law does not specify an exact penalty for a particular type of offense, but only a range of permitted penalties; the judge must then select some penalty within the permitted range which strikes him as appropriate, everything considered. The executive department of the government, however, has the power to reduce a sentence or pardon a criminal altogether for any reasons which strike it as proper; but it does not have the power to increase the penalty set by the judge. Various boards established by the executive department may pardon or parole, or make recommendations for such actions. Theoretically juries decide only a factual question: whether the accused acted in a way defined by the law as criminal; but practically they often refuse to convict, when they regard the penalty prescribed by law for a certain offense as markedly out of line with the moral merits of the case. In some instances juries may make recommendations about the severity of sentences.

In order to get a deeper understanding of what it is to commit a certain crime, it is helpful to draw an analogy between legal and moral concepts. Let us recall the relations between prima facie obligation, over-all obligation, and blameworthiness. In general we incline to think that, for instance, if a person has promised to do something, he has a prima facie obligation to do that thing. Sometimes, however, he will not have an over-all obligation to do this, on account of the weight of conflicting obligations, such as obligation to avoid serious injury to other persons, or to give important assistance in case of

need. A prima facie obligation to keep a promise, then, does not imply an over-all obligation to keep it. Moreover, failure to keep a promise, even if one is over all obligated to keep it, does not necessarily imply moral blameworthiness. . . .

Corresponding roughly to objective prima facie obligations are the law's prohibitions of certain kinds of conduct. The law aims to prevent certain kinds of overt behavior, and sometimes the law states, in a preamble preceding clauses prescribing punishment, what kind of conduct it is aiming to prevent. For instance, it may say that "No company shall sell, . . . or offer for sale . . . any security of its own issue until it shall have . . . secured . . . a permit authorizing it so to do." (California Corporate Securities Act, Cal. Stats. 1917, p. 673.) But the law recognizes that conduct of the sort it aims to prevent in general may sometimes be justified; and therefore it permits, as defense against a criminal charge, a showing that the accused's action was justified (that in the special circumstances the agent did what the law does not really want to prevent). For instance, the law forbids killing another person; but it is prepared to accept the defense that the killing was in self-defense. Roughly, a defense of justification is analogous to a showing, in morals, that although one had a prima facie obligation not to do so-and-so, in the total circumstances doing this was the right thing. Moreover, just as infraction of overall obligation does not imply blameworthiness in morals, so the commission of some act the law intends to prevent is not necessarily criminally culpable, even if it cannot be justified in the foregoing sense. Just as there may be excuses in morals for doing what one ought not, so there are excuses in law. It is useful to distinguish three kinds of excuses in law: (a) excuses which completely exculpate, wholly free from taint of crime, (b) excuses which mitigate in the sense of reducing the crime to one of the types which the law regards as less serious (e.g., reducing from murder in the first degree to involuntary manslaughter), and (c) considerations which are properly viewed by the judge as calling for the selection of a lower penalty from among those permitted by law.

We can now see more clearly what it is to have committed some specific crime, for instance, murder in the first degree. It is (1) to have behaved in some overt way which the law aims to prevent, here to have caused the death of someone, (2) to be unable to show a legally acceptable justification of one's act, and (3) to be unable to offer an excuse of the above types (a) or (b). In a carefully written penal code, the crime of murder in the first degree will be carefully described so that one will not have committed it unless all these conditions are satisfied. To commit a certain crime, then, is to behave in a certain way, and be unable to offer one of these defenses.

It will illuminate the relation between legal liability to punishment and moral excuses and blameworthiness, to review the major types of defense. We begin with the major *justifications.* (1) An act, otherwise subject to punishment, is not a crime if done in reasonable self-defense against unlawful attack, by a soldier in execution of lawful orders, in order to prevent treason or a felony (at least one done with violence), or in the service of public welfare (as destruction of property in order to prevent flood damage or to bring a forest fire under control). (2) An act, otherwise subject to punishment, is not a crime if performed because someone threatened the agent with loss of life or personal injury, in such circumstances as would intimidate a person of ordinary firmness. This justification, however, is ordinarily not accepted for homicide. (3) One may perform a forbidden act with impunity if, in case one does not, the result one causes and worse will in all probability happen anyway. For instance, one judge has indicated that a seaman is not liable for putting persons out of an overcrowded lifeboat in a storm if otherwise the boat would sink, provided the selection of the unfortunates is made in a reasonable manner. (For a contrary opinion, see *Regina v. Dudley and Stephens.*) This is one sense in which "necessity knows no law."

Let us now survey the *excuses which exculpate.* (4) A person is usually exculpated if he causes a harm the law seeks to prevent, by accident, for instance, if he causes the death of someone, but unforeseeably, despite all reasonable precautions. However, such accidental harm is sometimes punished; in many situations in which the defendant is committing an unlawful act when the accident happens, he is treated as if he intended to cause the accident. (5) A man is exculpated if because of an

innocent mistake in belief about facts he performs an unlawful act which would have been lawful had the facts been as he thought they were. For instance, it is lawful to shoot a person who has broken into one's house at night for the purpose of theft; and if a person shoots his wife or servant on account of a genuine and innocent supposition that the person fired at is a burglar, he is exonerated.[1] (6) If a person is physically compelled to perform an unlawful act, e.g., if the hands of a stronger person held and guided his dagger, he is thought not to have performed that act at all, and will not be prosecuted. (7) The law excuses a man for doing what it is impossible not to do. It may forbid a car standing in a certain place, but if a car stands there because traffic is jammed and the car cannot be moved, the act is not subject to punishment. (8) A person under seven years of age cannot commit a crime; and he is not guilty of a crime if under fourteen unless malicious intent is established—although he may be subject to treatment under rules for juveniles. (9) A person is excused if he performed an otherwise unlawful act while walking in his sleep. Insanity is also an excuse if the defendant's mental disease or defect is such that he did "not know the nature and quality of the act he was doing; or, if he did know it, that he did not know he was doing what was wrong." (*The Queen v. McNaghten.*)[2] (10) Theoretically a person is excused from guilt if he performs an unlawful act because of involuntary intoxication—even though in practice courts are often reluctant to admit that intoxication is "involuntary."

There are also *mitigating* excuses, which reduce the seriousness of the crime. (11) In the case of homicide, a showing that the act was done impulsively often affects the degree of guilt. Premeditation is commonly necessary for first degree murder. If one kills in the heat of passion, without forethought, he is guilty only of second degree murder. Furthermore, if such heat of passion is brought about by legally recognized provocation—such as finding one's wife in the act of adultery—the defendant may be guilty only of manslaughter. (12) Voluntary intoxication may reduce the degree of crime, but it will not exculpate, on the ground that intoxication is itself an immoral act which cannot be used as a shield for wrong conduct.

Finally, judges regularly take into account various considerations as grounds for reducing the severity of a sentence, such as severe temptation, neurotic constitution, previous lack of opportunities in life, evidence indicating that the culprit is not a menace to society, the convicted person's state of health, probable serious effects of a harsh sentence on his family. Judges sometimes make a sentence *more* severe when they think the outrage of the community requires it—if the law would be brought disrespect by a mild sentence.

The excuses (nos. four to twelve) are of interest because they give practical content to the "mental element" or "guilty mind" (*mens rea*) which the law holds must be present in order for a person to be guilty of a crime on account of his overt behavior. In view of these excuses, is there any general statement possible about the identity of this mental element? There are several things we can say. (1) It must have been possible for the accused to have behaved, as a result of different decisions or volitions, in a way that was lawful (a requirement of excuses nos. six and seven). (2) The agent must have been aware to some degree that the events which the law proscribes would flow from his bodily behavior—or at any rate that they might, that his behavior was *risking* such consequences. (3) The agent must have been aware that it is wrong to produce such consequences, *or* that it is unlawful, *or* at least that society generally regards the production of such consequences as wrong. In general, all these three conditions are required. There are two major exceptions: the so-called "public welfare" or "strict liability" offenses, where *intent* in the sense of (2) is not required at all; and the rule which imputes intent to commit offenses which in fact were accidental, to persons engaged in committing some other wrong intentionally. These exceptions are regarded by many as simply illogical elements in the law, which are without justification.

These conditions required for legal guilt are not exactly identical with those required for moral reprehensibility, the latter requiring that an act show *defect of character.* Moral guilt usually goes with legal guilt, but not necessarily. In practice, imputation of legal guilt is likely to be even more closely conformed to imputation of moral guilt than theory allows; for an unlawful act which is clearly consistent with moral character may not be prosecuted, or if prosecuted the jury may refuse to bring in a verdict of guilty.

How nearly is there a parallel between degree of moral reprehensibility of an act and seriousness of the act as a crime? We must expect some divergencies. The law, for instance, cannot concern itself with minor matters. Again, the law must be drawn in terms of relatively precise general rules, so that its administration does not place an intolerable burden of discretion on the officials charged with its application. Further, the law must be so phrased that available evidence can answer questions about its applicability; it is unfortunate if application of the law must turn on decisions about fact which are necessarily speculative.

At some points moral reprehensibility is more severe than the law. The law condemns only for overt behavior (including attempts) which is in some way publicly undesirable; it never condemns a man for his motives alone, or for his thoughts or feelings. Again, there are morally indefensible injuries of other persons (for instance, refusal to repay borrowed money) which, although they are actionable in civil suits, are not crimes and not punishable in criminal law. Again, there are many minor offenses that are morally wrong but legally are not prohibited. Further, the law automatically excuses offenses in persons under a certain age; moral judgments are more flexible and individual.

At other points, however, the law is more severe. (1) For instance, it is an axiom of the law that "ignorance of the law excuses nobody." And, we can add, it is no excuse to believe that it is morally right to do what the law forbids. Now, to some extent the same principles obtain in moral judgments: we would think it odd if anyone tried to excuse murder or rape on the ground that he did not believe the act to be wrong. Nevertheless, within a certain range, absence of belief that an act is wrong (or positive belief that it is a duty) serves as an excuse in morals; we hardly condemn a Mormon for practising polygamy, and we partly if not wholly excuse a Christian Scientist who refuses to take an ill child to the hospital. The law, however, does not regard the absence of belief that an act was unlawful as an excuse. Part of the reason for this is that the law in major part is enforcing moral rules that are well known and that are respected as right by the vast majority. (This class of cases, however, does not include many regulations that a moral person could hardly be expected to observe as a part of moral behavior, for example, peculiar traffic regulations in a small town. It is not clear that there is any moral justification for the law's refusing to accept ignorance as an excuse, in these cases where law fails to coincide with morals and where a man of character could hardly be expected to know the law.) There is also reason for not permitting conscientious objection to the law as an excuse: the law cannot permit a man to set up his own conscience as the law of the land, since if it did law enforcement could probably not proceed at all, or at least only under great handicaps. So, both the Mormon and the Christian Scientist may go to jail on serious charges.

An even more striking difference (2) between law and moral judgment is the existence of "strict liability." There is a class of infractions of law that are *not excused by demonstrable lack of intent;* any infraction of these laws is liable to punishment even if it was unforeseeable, inadvertent, and practically impossible to prevent. Thus, a butcher who sells diseased meat innocently and without any negligence, and even on the advice of experts about the quality of his meat, may incur a prison sentence. In general, a highly conscientious and law-abiding man might, through no moral fault whatever, infringe one of these laws and be subject to serious penalty. Moral condemnation certainly does not follow the law at this point.

Finally (3) some excuses that exculpate in morals only mitigate in law. In morals, whenever the circumstances are such that a man of character might be expected to do what the defendant did, he is exonerated. This is not so in law. Suppose a person imbibes a cocktail for the first time, not knowing that one drink will affect him so much that he will thoughtlessly accept another, and then be drunk. Suppose as a result of the first drink (and the second), he becomes drunk. Suppose further that this man has no reason to think that, even if he became drunk, he would be belligerent—beyond the vague awareness that some people do become belligerent when drunk. And suppose that, after becoming drunk, this person actually attacks a man and kills him. In law this can be a very serious offense; but in morals it is difficult to believe that the act is not blameless, as far as the homicide is concerned. The man *will* be morally to blame if, knowing what he did, he permits himself to

become drunk again. But his drunken action was not rationally foreseeable. (It is a different matter if a man drinks, knowing that he is going to drive a car immediately after doing so.) Similarly, the rule that homicide is automatically manslaughter if the agent was committing a misdemeanor, or murder in the first degree if he was committing a felony, is by no means followed by moral judgment; perhaps in most cases the assumption of intent is reasonable, but it is not necessarily so in all cases.[3]

It is often said that the law takes no account of *motives*, but only of the *intent* to perform a specific act. We have already noted that in practice this is not quite true. But we should also notice that, if we have a clear idea of what a *motive* is (namely, that for a person to have a certain motive is for him to be inclined or disinclined to perform a certain act by his belief that doing so will have a certain property or consequence), we can construe "intent" and *mens rea* as facts about a person's motives. Clearly, if a person intentionally performs a prohibited act, he thereby shows that he was *not motivated sufficiently* to avoid that act, and this fact is as much a fact about his motives as any other. What is true is that the law takes this fact about one's motives as very decisive for the legal culpability of one's act. But other motives also play an important role. One can, if one likes, construe justifying exculpations as based on a consideration of motives: for instance, one might say that the law recognizes the fact that the motive of saving oneself or one's children from death, or saving many from the destruction of a forest fire, is a good and sufficient justification for performing an unlawful act. Moreover, a showing of good motives or strong temptations may, although not exculpating from legal liability, result in a suspended or very light sentence (and, of course, may induce a jury not to convict at all—witness cases in which juries have refused to convict persons who practiced euthanasia under circumstances approved by the jury but against the law).

One might inquire whether any examination of the "ethical foundations" of the law can show that we are obligated to support and maintain the law at those points where it diverges from judgments of moral culpability. The answer is that it can, on matters where the law is less severe than morals (except for some of the artificial distinctions between crimes and torts—damages actionable in civil suits). It can because there is no alternative for law which is not excessively expensive and burdensome, or which would not introduce objectionably speculative elements into legal procedure. But, on some matters where the law is more severe than moral judgment, it cannot be shown that we ought to support the maintenance of present legal principles. It is quite true that (except for matters like traffic regulations where legal restrictions do not correspond with accepted judgments of moral obligation) it is impossible to accept ignorance of the law, or conscientious disagreement with the law, as an excuse. The law cannot cease to protect what it sets out to protect because some do not agree or are unaware. But the other practices of the law that morals cannot follow appear to be morally unjustified, and there is no obligation to defend them; on the contrary, there is an obligation to do what one can to get them changed.

NOTES

1. The supposition must be "innocent" in the sense that it is one a "reasonable" man might make in the circumstances. A person who supposes something negligently is not excused, except in some types of case.
2. The question of what should constitute legal insanity is the most vigorously discussed issue in criminal law today. For an important alternative view (in *Durham v. United States*) and discussion, see L. Hall and S. Glueck, *Cases on Criminal Law and Its Enforcement* (St. Paul, Minn.: West Publishing Co., 1958), pp. 298–325.
3. For a full discussion of some of these moral difficulties in legal practice, see Jerome Hall, *General Principles of Criminal Law* (Indianapolis: Bobbs-Merrill Co., Inc., 1947), chaps. 10, 11, and 13.

REVIEW AND DISCUSSION QUESTIONS

1. Give an example of a defense of justification. What is the difference between that and an excuse? How do exculpatory excuses, mitigating excuses, and considerations that lead judges to lower penalties differ?
2. Apply the distinctions that Brandt draws to the *Dudley and Stephens* case.

3. What is a "strict liability" offense, and how is it an exception to the general principles of criminal law?
4. In what ways might legal guilt and moral reprehensibility differ? At what points is morality severer than the law, and at what points is the law severer?
5. Explain what Brandt means when he writes that the common saying that the law takes no account of motive, but only of intent, is not quite true.
6. Are we obligated to support the law at those points where it diverges from judgments of moral culpability?
7. How does Brandt define *mens rea?*

Intention

H. L. A. Hart

Intention is the most prominent and in many ways the most important element of *mens rea,* and in this essay, Professor Hart examines the place of intention in the criminal law—with respect to both liability for punishment and severity of punishment. After describing three aspects of intention and the difference between oblique and direct intention, Hart discusses the role of intention in the law concerning attempted crimes and criminal negligence.

1

In this essay I shall [focus] attention on the place which the criminal law of most countries allocates to the idea of intention, as one of the principal determinants both of liability to punishment and of its severity. All civilized penal systems make liability to punishment for at any rate serious crime dependent not merely on the fact that the person to be punished has done the outward act of a crime, but on his having done it in a certain state or frame of mind or will. These mental or intellectual elements are many and various and are collected together in the terminology of English jurists under the simple sounding description of *mens rea,* a guilty mind. But the most prominent of these mental elements, and in many ways the most important, is a man's intention, and in English law and in most other legal systems intention, *or something like it,* is relevant at two different points. It is relevant first at the stage before conviction when the question is "Can this man be convicted of this crime?"—even if, in fact, he will not actually be punished. At this stage, so far as *mens rea* is concerned, it is normally, though not quite always, sufficient, and normally, though not quite always, necessary that the accused did the particular act forbidden by law, and did it intentionally or with something like intention. It is true that it is not always sufficient, because sometimes duress or provocation or certain forms of mental abnormality may become relevant, and in such cases the accused may not be convicted for a particular crime even if he intended to do the act forbidden by law. But the scope of these matters is small indeed; provocation, for example, is in English law limited to homicide, and duress does not extend to it. So intention, or something like it, is usually, though not quite always, *sufficient* so far as murder is concerned, for conviction of a man who has killed another. Intention to do the act forbidden by law, or something like it, is also generally *necessary* for serious crime, though there are exceptions. Gross unthinking negligence may be enough, e.g., in certain cases of manslaughter, as is carelessness in certain motoring offences, and there are also certain forms of strict liability when a man may be liable for punishment even though he did not intend to do what the law forbids, and was not even guilty of negligence. In regard to

murder, the law, in certain types of case has applied doctrines of "constructive" murder and "objective" tests of liability so as to render the question of the accused's actual intention largely irrelevant. But these are exceptions, and there is no doubt of the central importance of intention or something like it when the question is, "Is the accused liable for punishment?"

Intention is also relevant when the stage of conviction is past, and the question is, "How severely is the accused to be punished?" This is the stage of sentencing, as distinguished from conviction. Sometimes the legislature will mark off a greater maximum penalty for things done with a certain intention than for the same thing done without that intention. So wounding with intention to kill (even though the victim is not killed) or wounding with intention to resist arrest is punishable with the maximum penalty of life imprisonment,[1] and thus much more severely than the simple offence of "unlawfully and maliciously wounding" a man, for which the maximum penalty is five years.[2] Sometimes, however, the greater severity of punishment is settled not by the legislature, but is a matter for the judge to settle within the exercise of his discretion, and in doing this he may often allow the question of intention to weigh. However, too much importance should not be attached to these varying maximum penalties in the case of statutory crimes, since many different features may account for the variation. No philosophical principles, presumably, are needed or competent to account for the fact that, whereas maliciously damaging a work of art is punishable with a maximum of six months imprisonment, maliciously damaging textiles is punishable with imprisonment for life.[3]

2

The law's concern with intention at these two different stages (conviction and sentence) generates a number of problems, some of which I consider here. But there is first an analytical question to be faced before we reach these problems. What, after all, is a man's intention? This is a question which, quite apart from the law, philosophers have found both intriguing and enormously difficult to answer in any simple terms, and there is an additional difficulty in the case of the law. For though jurists or expositors of the criminal law will certainly speak as I have done up to now, and say that the notion of a man's intention is relevant to his criminal responsibility at many important points, what they refer to is the use by the law of a concept which, though it corresponds at many points to what is ordinarily meant in non-legal use by intention, cannot be said to be identical with it. . . .

What in the law corresponds to intention can . . . be made clear in the following way. Intention is to be divided into three related parts, to which I shall give what I hope are three self-explanatory names. The first I shall call "intentionally doing something"; the second "doing something with a further intention," and the third "bare intention" because it is the case of intending to do something in the future without doing anything to execute this intention now. The following are simple legal examples of these three aspects of intention. Suppose first that a man has done something which fits the definition of a crime so far as concerns the outward movements of his body which he has made and the harmful consequences; he has, for example, fired a gun at and thereby wounded or killed another man. On these facts the question then arises, "Did he wound (or kill) him intentionally?" Though these facts may, in the absence of further evidence, entitle others to conclude (both in and out of the law) that the answer to this question is "Yes," further evidence may show that he wounded or killed the other accidentally owing to a mistaken belief that the gun was unloaded, and so did it unintentionally. With such cases contrast those illustrating the second aspect of intention: doing something with a further intention. A man gets into a dwelling house at night and the question is not, or not merely, "Did he do that intentionally?" but "Did he do that with the further intention, or (as lawyers like to say) 'with the intent,' of stealing something?" If so he is guilty of burglary, even if in fact he did not steal anything. Many statutory crimes are framed in terms of such further intent, including the crime already mentioned of wounding with intent to kill, or with an intent to resist arrest, as contrasted with a simple "malicious" wounding which, subject to the considerations mentioned below, is equivalent

merely to intentionally wounding. The third aspect of intention, bare intention, or just intending to do something in the future without taking any present steps towards its execution, is not, for reasons which I shall mention later, of central importance in the criminal law, though it is important in the civil law. Indeed, a landlord's right to eject a tenant on the termination of a lease may depend on the question whether he intended before its expiration to reconstruct the premises. Of these three notions the first, intentionally doing something, is for legal purposes the most important, and I shall begin my discussion of problems by pointing to the divergence between the legal use of this notion and common usage. This shows itself in the following way. . . .

. . . Consider . . . the facts of a famous Victorian case, *R. v. Desmond, Barrett and Others.*[4] In 1868 there lay in jail two Irish Fenians, whom the accused attempted to liberate. For the purpose, one of them, Barrett, dynamited the prison wall outside the area where he mistakenly believed they would be at exercise. Though the plot failed, the explosion killed some persons living nearby. In this case we can distinguish Barrett's movements made in igniting the fuse from the harmful upshot or consequences in the death of the victims. Given such facts, we may ask outside the law, "Did he kill those men intentionally?" Inside the law the cognate question is "Did he kill them with such malice aforethought as is required to constitute murder?" Generally speaking, so far as any question like that of intentionally harming is concerned, the law, though it may also be content with less, is content to hold a man guilty if the harmful consequence, e.g., death, was foreseen by the accused in the sense that he believed that it would come about as a result of some voluntary action on his part. Whether he thought this would be certain or only likely to ensue may . . . determine the choice between the words "intentionally" or "recklessly" for the description of the case, but will not affect the accused's liability. But the point to be observed here is that, for the law, a foreseen outcome is enough, even if it was unwanted by the agent, even if he thought of it as an undesirable by-product of his activities, and in Desmond's case this is what the death of those killed by the explosion was. It was no part of Barrett's purpose or aim to kill or injure anyone; the victims' deaths were not a means to his end; to bring them about was not his reason or part of his reason for igniting the fuse, but he was convicted on the ground that he foresaw their death or serious injury. As Lord Coleridge said in Desmond's case, it is murder "if a man did [an] act not with the purpose of taking life but with the knowledge or belief that life was likely to be sacrificed by it."

The law therefore does not require in such cases that the outcome should have been something intended in the sense that the accused set out to achieve it, either as a means or an end, and here the law diverges from what is ordinarily meant by expressions like "he intentionally killed those men." For outside the law a merely foreseen, though unwanted, outcome is not usually considered as intended, and this is so in big matters as well as small. The neighbour who for the pleasure of the music plays her gramophone at 6 a.m., well knowing from my frequent complaints that it will wake me from sleep, as it does, would not normally be said to have intentionally woken me up or woken me up intentionally, any more than Barrett in ordinary parlance would have been said to have killed the victims of the explosion intentionally. . . . Some legal theorists, Bentham among them, have recorded this divergence by distinguishing (as "*oblique* intention") mere foresight of consequences from "*direct* intention" where the consequences must have been contemplated by the accused not merely as a foreseen outcome but as an end which he set out to achieve, or as a means to an end, and constituted at least part of his reason for doing what he did. . . .

It is perhaps easy to understand why when a man is accused of killing and the question is "Is he to be convicted and so liable to punishment for this?," as distinguished from "How severely shall we punish him?," the law should neglect the difference between oblique and direct intention, and why lawyers should come as they have in England to use the beautifully ambiguous expression, "He contemplated this outcome," to cover both. The reason is, I suggest, that both the case of direct intention and that of oblique intention share one feature which any system of assigning responsibility for conduct must always regard as of crucial importance. This can be seen if we compare the actual facts of the Desmond case with a case of direct intention. Suppose Barrett shot the prison guard in order to

obtain from them the keys to release the prisoners. Both in the actual Desmond case and in this imaginary variant, so far as Barrett had control over the alternative between the victims' dying or living, his choice tipped the balance; in both cases he had control over and may be considered to have chosen the outcome, since he consciously opted for the course leading to the victims' deaths. Whether he sought to achieve this as an end or a means to his end, or merely foresaw it as an unwelcome consequence of his intervention, is irrelevant at the stage of conviction where the question of control is crucial. However, when it comes to the question of sentence and the determination of the severity of punishment it may be (though I am not at all sure that this is in fact the case) that on both a retributive and a utilitarian theory of punishment the distinction between direct and oblique intention is relevant. . . .

3

. . . The law courts do apparently recognize this distinction between oblique and direct intention when they are confronted with that second aspect of intention which I have called further intention: that is, when the offence is defined as doing one thing with the intention that something else shall occur or be done, like wounding with intent to kill, or doing an act likely to assist the enemy with the intention of assisting the enemy. Here it seems that it is not enough that the accused believed that there would be a certain outcome; for conviction on such a charge it must be shown that he contemplated that outcome as an end or as a means to some end. The case most frequently cited in support of this view of the law is *R. v. Steane.*[5] The facts of this case were that Steane, who was resident in Germany on the outbreak of war, was threatened by the Gestapo that he and his wife would be beaten up unless he broadcast enemy propaganda in English. To save them, he did broadcast, and after the war was charged with the offence of doing an act likely to assist the enemy with the intention of assisting the enemy. He was, in fact, acquitted on the ground that unless it was shown that he broadcast in order to assist the enemy the charge was not made out, and his mere knowledge that it was likely to assist the enemy was not enough. The decision seems to me to be based on a correct interpretation of the statutory language; but the moral or policy justification for acquitting Steane on this ground [rather than duress] has seemed far from clear to many critics. . . .

I shall leave the topic of direct and oblique intention to consider the more general problems generated by the law relating to attempts. As everyone knows, a bare intention to commit a crime is not punishable by English law. This has been often repeated from the Bench since Lord Mansfield in 1784 said: "So long as an act rests in bare intention alone it is not punishable by our law." The reasons for this are perhaps not far to seek. Not only would it be a matter of extreme difficulty to ferret out those who were guilty of harbouring, but not executing, mere intentions to commit crimes, but the effort to do so would involve vast incursions into individual privacy and liberty. The Victorian judge, James Fitzjames Stephen, said that to punish bare intention "would be utterly intolerable: all mankind would be criminals, and most of their lives would be passed in trying and punishing each other for offences which could never be proved."[6] But the law, though it does not punish bare intention, does punish as an attempt the doing of something quite harmless in itself, if it is done with the further intention of committing a crime and if the relationship between the act done and the crime is sufficiently "proximate" or close. So the would-be thief who puts his hand into a pocket which contrary to his expectations proves empty, or who writes a letter to obtain money by false pretences which fails to deceive his intended victim, or the would-be murderer who puts poison into the cup which is emptied before the intended victim can drink out of it, are all guilty of attempts to commit crimes.

It is not obvious, however, at least on some versions of utilitarian theory of punishment, why attempts should be punishable, as they are, in most legal systems. On a retributive view perhaps the answer is easy. The criminal had gone so far as to do his best to execute a wicked intention, and the difficulties of proof and so on are removed by his overt act. But on what is generally known as a deterrent theory the case for punishing attempts has seemed, even to some of its supporters, unclear. Thus it has

been argued that if we think of the law, as the deterrent theory requires us to think of it, as threatening punishment to those persons who are tempted to commit offenses, there can be no need to attach any punishment to the unsuccessful attempt, because those who set about crime intend to succeed and the law's threat has all the deterrent force it can have if it is attached to the crime; no additional effect is given to it if unsuccessful attempts are also punished. . . . [But] there seems a clear case for the use of punishment as an individual deterrent in the cases of unsuccessful attempts to commit crimes; for the accused has manifested a dangerous disposition to do all he can to commit a crime, and the experience of punishment may check him in the future, since it may cause him to attach more weight to the law's threats. From this point of view the punishment of a man who has attempted but failed seems as well justified on deterrent grounds as the punishment of a man who has succeeded in committing a crime, although I shall consider later the usual practice of punishing the attempt less severely. But even from the point of view of the general deterrent, the skeptical argument which suggested that there is no case for punishing an attempt is, after all, mistaken. It is perfectly true that those who commit crimes intend to succeed, but this does not show that punishing a man for an unsuccessful attempt will not increase the efficacy of the law's threats, or that failure to punish him would not often diminish their efficacy. This is so for two reasons: first, there must be many who are not completely confident that they will succeed in their criminal objective, but will be prepared to run the risk of punishment if they can be assured that they have to pay nothing for attempts which fail; whereas if unsuccessful attempts were also punished the price might appear to them to be too high. Again, there must be many cases where men might with good or bad reason believe that if they succeed in committing some crime they will escape, but if they fail they may be caught. Treason is only the most obvious of such cases, and unless attempts were punished, there would, in such cases, be no deterrent force in the law's threat attached to the main crime.

A more difficult question concerns the almost universal practice of legal systems of fixing a more severe punishment for the completed crime than for the mere attempt. How is this to be justified? Here a retributive theory in which severity of punishment is proportioned to the allegedly evil intentions of the criminal is in grave difficulty; for there seems to be no difference in wickedness, though there may be in skill, between the successful and the unsuccessful attempt in this respect. Very often an unsuccessful attempt is merely the accidental failure to commit the crime because somebody unexpectedly intervenes and frustrates the attempt. As far as I can see a deterrent theory, except in relation to a very specialized class of crimes, is in similar difficulties. . . . There seems no reason on any form of deterrent theory, whether we consider the general deterrent or the individual deterrent, for punishing the unsuccessful attempt less severely than the completed crime. The individual who has tried but failed to carry out the planned crime may need just as much punishment to keep him straight in the future as the successful criminal. He may be as much disposed to repeat his crime.

The almost universal tendency in punishing to discriminate between attempts and completed crimes rests, I think, on a version of the retributive theory which has permeated certain branches of English law, and yet has on occasion been stigmatized even by English judges as illogical. This is the simple theory that it is a perfectly legitimate ground to grade punishments according to the amount of harm actually done, whether this was intended or not; "if he has done the harm he must pay for it, but if he has not done it he should pay less." To many people such a theory of punishment seems to confuse punishment with compensation, the amount of which should indeed be fixed in relation to harm done. Even if punishment and compensation were not distinguished in primitive law, many think that this is no excuse for confusing them now. Why should the accidental fact that an intended harmful outcome has not occurred be a ground for punishing less a criminal who may be equally dangerous and equally wicked? I may be wrong in thinking that there is so little to be said for this form of retributive theory. It is certainly popular, and the nearest to a rational defense that I know of it is the following. It is pointed out that in some cases the successful completion of a crime may be a source of gratification, and, in the case of

theft, of actual gain, and in such cases to punish the successful criminal more severely may be one way of depriving him of these illicit satisfactions which the unsuccessful have never had. This argument, which certainly has some attraction where the successful criminal has hidden loot to enjoy on emerging from prison, would be an interesting addition to theories of punishment of the principle that the wicked should not be allowed to profit by their crimes.

My own belief is that this form of retributive theory appeals to something with deeper instinctive roots than the last mentioned principle. Certainly the resentment felt by a victim actually injured is normally much greater than that felt by the intended victim who has escaped harm because an attempted crime has failed. Bishop Butler, in his sermon on resentment, explains on this ground the distinction men draw between "an injury done" and one "which, though designed, was prevented, in cases where the guilt is perhaps the same." But again the question arises, if this form of retributive theory depends on the connexion between blame and resentment, whether the law should give effect to such a theory. Can we not control resentment, however natural, in the interests of some deliberate forward-looking policy, much as we control our natural fears in the interest of forward-looking prudential aims? And if we can do this, should we not do so? And might not this require us in some cases to punish attempts as severely as the completed offence?

Now to my last topic. I have said that the intention, as the law understands it (that is either oblique or direct intention), is generally, though not always sufficient and generally necessary for criminal liability. But it is not always so, and in conclusion I shall consider certain cases where a man may be punished for a crime, although he had no intention, oblique or direct, to do the act forbidden by law. Here too, we shall be forced in the end, if we are to fit any theories of punishment to the facts, to ask what we mean precisely by the idea of retribution or deterrence, and to refine these ideas in further new directions.

There are two main types of case where no form of intention is required for criminal liability. The first type consists of those crimes known as crimes of strict liability, where it is no defence to prove that you did not intend to do the act forbidden, did not know that you were doing it, and indeed took every care to avoid doing it. Strict liability of this sort is usually thought odious by academic writers, even though most of the offences in question are punished with fairly minor penalties. They include such things as selling liquor to an intoxicated person, or selling adulterated milk, or driving a car without insurance; though they also include more serious offences. The second type of case on which I shall spend a little time are cases of negligence, and I shall concentrate on this, since some useful lessons can be learnt from it.

English law is rather sparing in its punishment of negligence in the sense of unintentional neglect to take reasonable precautions against harm to others, and apart from a few isolated situations, it can be said that negligence is only punishable as a crime if it results in death (where it constitutes one species of manslaughter) or if it is shown in driving motor vehicles on the road: not only is there the offence of driving without due care and attention but there is also the more severely punished offence of causing death by dangerous driving. There are no doubt some general historical reasons why the law should be slow to punish a man who is negligent in the sense of not realizing, though he ought to have realized, that his actions or omissions might occasion a serious harm; but in addition to these historical grounds it is the case that many lawyers feel uncomfortable in accepting negligence as a basis of criminal liability, though, in fact, in our ordinary life it is not usually held to be a good excuse to say "I didn't mean to do it" or "I just didn't stop to think." Certain jurists draw a very dramatic line between the cases where a man intended or foresaw that his actions would be harmful, and those where he was grossly careless, but, as they say, "inadvertent." To such lawyers it appears that the "state of mind" of the merely careless man is not in any way wicked, and presents nothing on which a retributive theory can, as it were, bite. But what is more extraordinary is that very many utilitarian-minded jurists think that there is no intelligible case for punishing gross inadvertent negligence as a deterrent. This seems to conflict very much with the common-sense belief that in some cases we may make people more careful by blaming or punishing them for carelessness. . . .

Threats may not only guide your deliberations—your practical thinking—but may cause you to think. A man drives a car with one arm round his girl friend's neck and gazes into her eyes instead of at the road, and is subsequently punished for careless driving in spite of his protestations that he "just didn't think" of the possibilities of the harmful outcome to others on the road. Surely it is not absurd to hope that, as a result, next time he drives he may approach his car, and perhaps his girl friend, in a very different spirit. Recollection of the punishment and the knowledge that others are punished may make a driver think; and if he thinks (since he has no intention to drive badly) he may say to himself "this time I must attend to my driving," and the effect may well be that he drives with due care and attention. No doubt the connexion between the threat of punishment and subsequent good behaviour is not of the rationalistic kind pictured in the guiding-type of case. The threat of punishment is something which causes him to exert his faculties, rather than something which enters as a reason for conforming to the law when he is deliberating whether to break it or not. It is perhaps more like a goad than a guide. But there seems to me to be nothing disreputable in allowing the law to function in this way, and it is arguable that it functions in this way rather than in the rationalistic way more frequently than is generally allowed. At any rate, consideration of the punishment of negligence (and also punishment in the strict liability cases) brings out the need to refine in this way the idea of deterrence by threat.

The punishment of negligence has in England, and I suspect in most legal systems, some further curious features, for here the severity of punishment is often determined by the seriousness of the outcome, and the explanation of this involves recourse again to that sense of retribution which I have already mentioned in explaining the lighter punishments normally accorded to unsuccessful attempts to commit crimes. In 1956, when the offence known as "causing death by dangerous driving" was created, there was an illuminating debate in the House of Lords, where Lord Hailsham pointed out how absurd it was that, where two people were equally careless in driving on the roads, one of them should be liable to be punished with the severe sentence of five years' imprisonment if the bad driving resulted in someone's death, whereas if it resulted only in the victim being crippled, or if no one was harmed he would be liable only for the maximum penalty of two years. Of course, a similar "illogicality" as Lord Hailsham called it, is to be seen in the very existence of one species of manslaughter where the accused's grossly negligent act is punishableif it causes the victim's death. The then Lord Chancellor (Lord Kilmuir) recognized the "illogicality" of this mode of determining the relevant severity of punishment by reference to the harm done, but said in the debate, "Such doubts have of course affected every penal thinker and penal reformer. But no one has been able to translate these doubts into a workable system. Results must be taken into account if the penalties are going to have the effects which it is desirable they should have." I have never understood that answer in defence of this form of retributive theory. Does it merely state obscurely what Stephen stated with great clarity seventy years earlier, "It gratifies a natural public feeling to choose out for punishment the one who actually has caused great harm"?[7]

NOTES

1. Offences against the Person Act. 1861, s. 18.
2. Ibid., s. 20.
3. Malicious Damage Act. 1861, ss. 39 and 14.
4. *The Times*, 28 April 1868 (hereinafter referred to as "Desmond's case").
5. (1947) K.B. 997.
6. Stephen, *A History of the Criminal Law of England*, Vol. II, p. 78.
7. Stephen, op. cit., Vol. III, p. 311.

REVIEW AND DISCUSSION QUESTIONS

1. What are the two different stages at which the law is concerned with a defendant's intention?
2. Explain the three aspects of intention that Hart distinguishes.

3. What principle does the *Desmond* case illustrate?
4. What is the difference between oblique and direct intention? Why does Hart believe that in cases of killing the distinction is relevant to the severity of punishment but not to the question of liability to punishment?
5. What did the Court decide in *R. v. Steane?* Was the decision a sensible one?
6. Should attempts be punished less severely than successful crimes? Why or why not?
7. When, if ever, should negligence be criminally punished? Explain your answer. Is the seriousness of the outcome relevant to sentencing? Why or why not?

Attempting the Impossible: The Crime That Never Was

Leo Katz

Although criminal liability for attempting to commit a crime is relatively recent, it is now well established in the criminal law and for good reasons, both retributivist and utilitarian. But difficult questions arise in cases in which it is literally impossible for a person to have successfully committed the crime he or she is accused of attempting to commit. In addressing such cases, the courts have ruled in different ways, and it is difficult to determine what principles should guide judges in their decisions. In this essay, Leo Katz offers an imaginary case that highlights the puzzles that can arise when it is impossible for a person to succeed in committing the crime in question. Although it is, in Katz's words, "a fake opinion in a fake case involving fakes," the precedents to which his fictional judges refer are real.

COMMONWEALTH V. OMEIRA SUPREME COURT OF WESSEX

Before NEWSON, C. J., HENCHARD, HARDY, FARFRAE, Middlebury, J. NEWSON, C. J.: The defendant, Jan Omeira, is charged with attempting to export illegally "valuable artifacts of the native culture" in violation of Section 901.34(1) of the Wessex Code.

To protect our national patrimony, the Wessex legislature twenty years ago passed a law making it illegal to export native art works produced before 1920 and worth more than 100,000 pounds. This sweeping prohibition is subject to only one ill-defined exception. The Arts Council, an agency of the Interior Department, has broad discretion to grant exemptions, that is, special licenses, when it finds that "unusual circumstances so warrant." Licenses are typically granted when a work of art cannot be sold in Wessex except at a very small fraction of its world market price or when the applicant promises to secure another "native" work of comparable worth currently in foreign hands.

The defendant is a retired businessman of considerable wealth. He owns a farm in Casterbridge, Wessex, where he spends half of the year, and a rancho in Cuernavaca, Mexico, where he spends the other half. He is a zealous aficionado of the arts; indeed, since his retirement the acquisition of expensive paintings has become his chief preoccupation. He has opened art galleries both in his house in Casterbridge and in his villa in Cuernavaca, the one in Casterbridge for the public at large, the one in Cuernavaca almost exclusively for his own use and that of his guests.

The Constabulary has been aware for some time that not all native paintings Mr. Omeira is known to have acquired at public auctions in Wessex have made their appearance in his Casterbridge gallery and suspected him therefore of having illegally transported many of them to his

gallery in Cuernavaca. Since few people have gained access to that gallery, no direct confirmation has been possible. For several years Mr. Omeira was systematically strip-searched on every one of his departures for Mexico. The searches turned up nothing and were discontinued.

In January of this year, the Constabulary received an anonymous tip that some time in March, Omeira would personally be smuggling a picture by the famous eighteenth-century landscape painter Ignacius Decameron out of the country. The picture, *Seminole Falls,* had for several years been hanging in Omeira's Casterbridge gallery and had recently been taken down for "cleaning." When Omeira left for Mexico in March, Customs scoured his luggage. They found a false bottom in one of his trunks and inside it, tightly rolled up, *Seminole Falls.*

Omeira's case was about to go to trial, when something very unusual happened. A well-known Belgian landscape painter named Flammarion remarked in an interview with the French monthly Paris Match that he thought it "rather amusing" that a Wessex art collector was being prosecuted for smuggling a Decameron landscape when in fact the painting was a forgery. How did he know it was a forgery, the astonished journalist asked. Because, replied Flammarion, he had painted it himself.

The government to whom Omeira had forfeited the smuggled picture immediately invited a panel of experts to test Flammarion's claim. They found that the painting's age crackle had a different structure from that of genuine eighteenth-century paintings. Moreover, the crackle seemed to have been artificially produced. They found that the dirt in the painting's crevices wasn't dirt at all. It had crept into the crevices very unevenly and had a different homogeneity from dirt. They suspected it was ink. With the help of radio-graphic studies they discovered a residue of a prior painting underneath the landscape, suggesting that this was an eighteenth-century canvas that had been recycled for forgery. They also noticed some of the paints used contained pigments not known in Decameron's day. . . . The group concluded that Flammarion was right: the picture was a fake.

. . . For nearly ten years [before Omeira purchased it] the painting remained a part of the VanDamm collection. Then VanDamm, who found himself in financial straits, decided to sell it. He applied to the Arts Council for an export license so that he would be able to offer it up through Sotheby's. An expert from the Arts Council inspected the painting and concluded that it was probably a fake, worth at best 50,000 pounds. Although under the circumstances no license would have been necessary, the council issued one anyway, simply because the painting's status was still unclear.

VanDamm, however, had lost all appetite for having it sold at Sotheby's. If the painting really was a forgery, a public auction would bring that to light all too quickly. Instead, he discreetly searched for buyers among his colleagues in the business. He finally sold it to Jan Omeira for 1,200,000 pounds. Needless to say, he mentioned nothing about the painting's suspect provenance or even about the export license.

The discovery that *Seminole Falls* was a forgery put the trial judge in this case in a delicate quandary. Omeira was charged with attempting illegally to export a valuable native painting. But it is perfectly legal to export a forged Decameron. How then was the defendant guilty of any wrongdoing? The trial judge chose to slight the issue. In a disturbingly desultory opinion, he simply noted that the "defendant believed he was smuggling a real Decameron. Therefore he is guilty of attempting to export a 'valuable artifact of native culture' " and sentenced the defendant to three years in prison. The defendant appealed. I believe his appeal has merit.

A failed crime can still be a crime. That's why we have the law of attempts. The assassin who is prevented from firing a bullet by an alert bodyguard, the safecracker who is stopped short of opening the vault by an unsuspected alarm system, the rapist who is frustrated in his aim by an obstreperous victim, all have failed in completing their intended crime; yet they are guilty of a crime nonetheless, the crime of criminal attempt.

But not all failed crimes are crimes. "Suppose a man takes away an umbrella from a stand with intent to steal it, believing it not to be his own, but it turns out to be his own, could he be convicted of attempting to steal?" Baron Bramwell, who posed this hypothetical more than a century ago in a case called *Regina v. Collins,* rightly considered the question purely rhetorical.

When is a failed crime not an attempt? It behooves us to make a brief foray through some of the more typical cases and to see what general principle is to be extracted from them.

In *Commonwealth v. Dunaway* a man was charged with attempting to rape and engage in incestuous relations with his daughter. He had apparently advanced quite far in this undertaking when his wife called the police to arrest him. In the course of the trial it turned out that the girl was the man's stepdaughter. Wessex law makes consanguinity a prerequisite of incest. The defendant did not, of course, know that; he thought he was committing incest. He was convicted. On appeal the attempted rape conviction was upheld, the attempted incest conviction overturned. The court observed that even if the man had succeeded in his undertaking, even if he had actually completed an act of intercourse with his stepdaughter, he would only be guilty of rape, not incest. Stephen's Digest of the Criminal Law defines an attempt to commit a crime as "an act done with intent to commit that crime, and forming part of a series of acts which would constitute its actual commission if it were not interrupted." The defendant's act, even if not interrupted, would not have constituted incest. Hence, the court reasoned, he could not be guilty of the attempt to commit incest.

The defendant in *Stephens v. Abrabams* wanted to import a certain item into Victoria (Australia) without paying the duty on it. To this end he presented the customs officer with a fake invoice for the item. Unbeknownst to him, the item was not dutiable anyway. A bill was pending in Parliament that proposed to tax such items, but it had not yet been passed. The Customs Office discovered the defendant's deception and charged him with attempt to "defraud the revenue contrary to the Commonwealth Customs Act." The Supreme Court acquitted him. Even "if the accused had succeeded in his object, he would not have succeeded in defrauding the revenue," it argued. Even if the defendant had managed to deceive the Customs office with his fake invoice, he would not have cheated them out of any money they were entitled to. So there could be no attempt to defraud them either. . . .

In *People v. Dlugash* the defendant was charged with attempted murder because he had shot a corpse.

> Defendant stated that, on the night of December 21, 1973, he, Bush and Geller had been out drinking. Bush had been staying at Geller's apartment and, during the course of the evening, Geller several times demanded that Bush pay $100 towards the rent on the apartment. According to defendant, Bush rejected these demands, telling Geller that "you better shut up or you're going to get a bullet." All three returned to Geller's apartment at approximately midnight, took seats in the bedroom, and continued to drink until sometime between 3:00 and 3:30 in the morning. When Geller again pressed his demand for rent money. Bush drew his .38 caliber pistol, aimed it at Geller and fired three times. Geller fell to the floor.

Then, to confuse the police and to buy the defendant's silence, Bush ordered him to fire some extra bullets into Geller's body. The somewhat frightened defendant did just that, believing that Geller was still alive. The autopsy, however, revealed that Geller was almost certainly already dead. In an arcanely reasoned opinion, the New York Court of Appeals somehow reached the conclusion that the defendant could indeed be found guilty of attempted murder. Few courts, I venture to say, would accept that conclusion. Lord Reid in *Haughton v. Smith* some years ago contemplated just such a case and gave what I think is the definitive answer: "A man lies dead. His enemy comes along and thinks he is asleep, so he stabs the corpse. The theory [advanced by some] inevitably requires us to hold that the enemy has attempted to murder the dead man. The law may sometimes be an ass but it cannot be so asinine as that."

The defendant in *People v. Jaffe* had bought what he thought were stolen goods from some undercover policemen. He could not be charged with buying stolen goods since the goods weren't stolen. Instead, he was charged with attempting to buy stolen goods. The courts acquitted him: "If all which an accused person intends to do would if done constitute no crime it cannot be a crime to attempt to do with the same purpose a part of the thing intended." . . .

These cases establish a simple yet powerful principle. An act which, unless interrupted, constitutes a crime is a criminal attempt. But an act which, even if completed, wouldn't be a crime, is not.

The defendant in all of these cases is morally heinous. But why is he morally heinous? Because

of what he did? No, because of what he thought he did. In that case to convict him "would be to convict him not for what he did but simply because he had a guilty intention." It is a fundamental tenet of our criminal law that a man cannot be convicted for his thoughts, only for his acts.

My disposition of Omeira's case should now be clear. Even if Jan Omeira had succeeded in smuggling his painting across the border, he would not have violated the export ban, since *Seminole Falls* is a forgery. If the completed act was no crime, the attempt could not possibly be one either. The trial court's verdict should be reversed.

Henchard, J.: I disagree sharply with the reasoning of Justice Newson. He would have us endorse what I think is a rather strange principle: that an "act which, although intended to be a crime, would not have amounted to one, even if it had not been interrupted, is not a criminal attempt." Taken seriously, this principle would have absurd implications. Justice Newson concedes that the assassin who is overpowered by a bodyguard, the safecracker who is caught by an unsuspected alarm system, the rapist who is stymied by a resistant victim all are guilty of a criminal attempt. But suppose that given the way the assassin aimed his gun, he would have missed his target anyway. Justice Newson's principle would have the man acquitted. Suppose the gold bars in the vault were too heavy for the safecracker to move, even if he had gotten to them. Justice Newson's principle would have the man acquitted. Suppose the rapist was impotent and could not have achieved an erection. Justice Newson's principle would have the man acquitted. I do not see how Justice Newson can propose a principle with such consequences.

Nor can I approve of many of the decisions that seem to have endorsed this principle. If the defendant in Jaffe thought he was buying stolen goods, he was attempting to buy stolen goods. . . . By contrast, the decision of the *Dlugash* court was exactly right: If the defendant in *Dlugash* thought he was shooting a human being, he was attempting murder. In each of these cases, what the defendants were attempting was impossible only because of some unforeseen contingency. That makes them no different from the assassin, the safe-cracker, the rapist. They, too, failed because some unforeseen contingency stopped them in their tracks.

Justice Newson contends that to punish the defendants in cases like *Dlugash* [and] *Jaffe* . . . is to punish evil thoughts, not evil acts. But that's not so. We punish the defendants there not because they wanted to commit an evil act but because they took what they thought were substantial steps toward putting those thoughts into practice. I think I can pin-point the source of Justice Newson's confusion. He thinks that when but for the defendant's evil thought he would not be punished we are punishing him for the evil thought. That's a mistaken idea. We would not convict a murderer but for the fact that he intended to kill a human being. Yet it can hardly be said that we are punishing him only for his evil thoughts.

I do not think, therefore, that Jan Omeira is innocent of a criminal attempt to violate the export ban on art merely because what he took to be a real Decameron turned out to be a forgery. But I think there are other reasons for acquitting him.

Suppose two men furtively engage in homosexual intercourse thinking that it is illegal. In fact, state law has nothing against mutually consented-to homosexual intercourse. Are they guilty of a criminal attempt? Evidently not. You cannot invent the law against yourself. Just because you think something is illegal and then attempt to do it, you haven't yet done anything illegal. The crime you attempt is, we might say, "legally impossible" because there isn't such a crime. This is very different from the case where what you are attempting to do isn't really criminal because the facts, not the law, are different from what you took them to be: because the man you attempt to shoot is already dead, because the goods you attempt to buy are not really stolen, because the method you adopt for killing someone won't really work. We might call these cases of "factual impossibility." In sum, attempting the legally impossible is not a crime, attempting the factually impossible is. In a way, this is a corollary to the principle that ignorance of law is no excuse. Just as thinking something is legal when it isn't won't get a defendant out of a bind, thinking it is illegal when it isn't won't get him into one.

This principle, rather than the one endorsed by Justice Newson, serves to make sense of the three cases cited in his opinion with which I agree. I agree that the defendant in *Dunaway* who thought he was committing incest should have been acquitted of

the charge of attempting incest. What he attempted was legally impossible. The law does not make intercourse with one's stepdaughter part of incest. Thinking that the law does cannot make the defendant guilty of attempted incest.

I also agree that the defendant in *Stephens v. Abrahams* who thought he was smuggling a dutiable item past customs should have been acquitted of the charge of attempting to "defraud the revenue." He, too, attempted the legally impossible, since Victoria did not make the item he smuggled dutiable. Thinking that it did could not have made the defendant guilty of attempted smuggling. . . .

When the Decameron was still in Mr. VanDamm's possession, he applied for an export license to the Arts Council. The Arts Council granted him the license. In effect, they amended the export law so as to exempt this particular painting from its sway. Thus what Mr. Omeira attempted to do was to smuggle out of the country a painting under the mistaken belief that the law prohibited him from exporting it. In fact, the law specifically exempted that painting. What he was attempting to do was not merely factually impossible (because he was dealing with a forgery) but legally impossible (because the painting had been exempted from the export ban). The case is thus on all fours with *Dunaway* and *Stephens*. . . . It is for that reason that I too would acquit Mr. Omeira, notwithstanding my wholehearted disagreement with the reasoning of Justice Newson's opinion.

FARFRAE, J., with whom HARDY, J., concurs: . . . I agree in spirit with Justice Newson's approach. I agree with him that many of the so-called impossible attempt cases should be resolved in the defendant's favor. I disagree with the particulars of his argument, for many of the reasons given in Justice Henchard's opinion. And I disagree with his resolution of this case.

I disagree both in spirit and substance with Justice Henchard's approach, and, of course, with his resolution of this case in particular. The approach hinges on a distinction that strikes me as both obscure and unimportant, that between law and fact, legal impossibility and factual impossibility.

The distinction between "law" and "fact" has proved obscure wherever it is employed. For instance, the common law used to require that a plaintiff's complaint in a civil action only state the "facts" of his case, not any "legal conclusions." Unfortunately, no one had ever been able to tell whether the allegation that "on November 9, the defendant negligently ran over the plaintiff with his car at the intersection of State Street and Chestnut Street" is a statement of fact or a legal conclusion. In fact, the distinction between law and fact is just the legal version of the philosophical distinction between "empirical" and "analytical" statements, a distinction on whose existence philosophers have been unable to agree to this day.

The distinction is as unimportant as it is obscure. It distinguishes between cases that are really alike. . . . The present case shows neatly just how unimportant it is. The defendant Omeira made two mistakes. First, he mistakenly thought the picture was authentic. Second, he mistakenly thought it hadn't been licensed for export. Under Justice Henchard's rule, the first mistake fails to exonerate him, but the second mistake does. The first is a mistake of fact, the second a mistake of law. Yet I fail to see any profound difference between the two kinds of mistake. If Omeira's ignorance of the painting's authenticity doesn't exonerate him, then neither should his ignorance of the export license.

The proper way to approach cases like the present is to ask two questions. Let me ask them in turn, to explain why they are important, show how one goes about answering them, and answer them for the present case. The first question is this: Did the defendant really attempt something criminal?

An attempt is often mistakenly thought of as the fragment of a completed offense. Of course, that isn't so. The driver who hurtles down a slippery road at breakneck speed may be inviting a deadly accident that would qualify as involuntary manslaughter if it occurred. Nevertheless, he is not attempting to commit the crime of involuntary manslaughter. To attempt something one must not merely be on one's way to committing it, one must intend to commit it. The reckless driver clearly is not.

This obvious point has subtle implications. It means that a defendant may be thinking he is committing a crime, without actually attempting

to commit it. He knows the bomb he plans to hurl into the queen's carriage will kill not only the queen, but her bodyguard, but he is only attempting to kill the queen, not the bodyguard. Killing the bodyguard is an unintended by-product of killing the queen. Of course, determining whether somebody is actually attempting something or merely engaging in conduct which he thinks will bring it about, often is hard. It depends on whether the commission of the crime is his desired end or a means toward such an end (in which case we have an attempt) or whether it is rather a by-product of bringing about some desired end. To find out which it is, one has to ask whether the defendant would change his course of conduct if he thought the commission of the crime was not tied to the achievement of his desired end.

If we apply this analysis, we will see that many defendants charged with impossible attempts are not in fact attempting the crime they are charged with attempting. They merely think they are committing a crime. The rapist in *Commonwealth v. Dunaway* is not guilty of attempted incest, because he was not intending to commit incest. He only thought he was committing incest. Had he been told that incest requires consanguinity, he would have been relieved. He would certainly not have desisted from his actions. The "smuggler" in *Stephens v. Abrahams* is not guilty of attempted smuggling because he was not intending to smuggle. He merely thought he was smuggling. Had he been told that the items he was importing weren't dutiable, he would have been relieved. He would certainly not have abstained from importing them. The "killer" in *People v. Dlugash* is not guilty of attempting murder, because he was not intending to kill the already-dead man. He merely thought he was killing him. Had he been told that he was shooting a corpse, he would have been relieved. He would certainly not have avoided shooting it. The "fence" in *People v. Jaffe* is not guilty of attempting to buy stolen goods because he was not intending to buy stolen goods. He merely thought he was buying stolen goods. Had he been told that the goods were not stolen, he would have been relieved. He would certainly not have eschewed buying them.

What now of this case? Was the defendant intending to export an authentic Decameron? Or was he merely thinking he was exporting an authentic Decameron? That depends: Had he been told that the Decameron was fake, would he have cared? Would he have changed his course of conduct? Would he not have exported it? The defendant will, of course, argue that although he thought the painting was a genuine Decameron, that was not the reason he wanted to export it. He will argue that he liked the painting for its artistic merits, not its provenance, and that he would still have wanted to take it to Cuernavaca, even if it was a forgery. That's a tough argument to reckon with.

The record indicates that the defendant liked to keep his most exclusive and prized possessions in his Cuernavaca gallery. It also indicates that the defendant did not collect art as an investment. Very few paintings he acquired he ever resold. He collected them purely and simply for the aesthetic pleasure they afforded. Asking whether the defendant would have tried to export a forged Decameron to Cuernavaca amounts to asking whether his aesthetic enjoyment of the painting would have been diminished by his discovery that it was fake. Should it have been? Is it rational to enjoy a painting as long as you think it is a Decameron and on learning it is a mere Flammarion-imitating-Decameron cease to do so?

. . . I don't think there is anything snobbish about such an attitude. A man is not irrational or unreasonable for behaving in this fashion. And I don't feel we are imputing any irrational, unreasonable or implausible trait to Mr. Omeira when we assume that he cared very much that his picture be a real Decameron rather than merely "another pretty picture."

We have established that the defendant really attempted to export an authentic Decameron. But that is not enough to show that he is guilty of a criminal attempt. Before finding him liable for that, we need to answer a second question: Did the defendant really create an unreasonable risk of a crime being committed? Why do I think this question needs to be asked?

Before we convict someone for recklessly or negligently causing harm, we require that his conduct be "unreasonable," that it be the sort

a reasonable man would take exception to. In a sense, that introduces an element of luck into the law. The defendant may think that what he is doing creates an unreasonable risk, but if in fact it does not, he will not be convicted. The law does not want to trouble itself with conduct that wouldn't bother a reasonable man.

We should impose the same requirement of unreasonableness before we convict someone for intentionally or knowingly causing harm. Indeed, I think we already do. It's just that, typically, when a defendant intentionally brings about harm, there is no doubt that his conduct was such as would have bothered a reasonable man. If it wouldn't have bothered a reasonable man, the prosecutor usually decides not to press charges. That has the unfortunate effect of making us overlook this potentially important point.

Let me elucidate with an example. Suppose a father wants to kill his five-year-old son. He decides to do so by sending him to summer camp, not, as he did in past years, by train, but by plane instead. He is under the mistaken impression that plane crashes are a lot more frequent than train crashes. He hopes such a crash will occur. And indeed it does. Clearly, the father has intentionally caused his son's death. (He intended his son to die in just the manner he did, and the son wouldn't have died, if he hadn't been on that plane.) But should we convict him of murder? I don't think so. Why? Because the father did nothing a reasonable man would object to: It was not unreasonably risky to send the boy to summer camp by airplane.

Suppose the plane never crashed. But the police learn of the father's evil intentions and charge him with attempted murder. Should we convict him? I don't think so. Why? Because the father did nothing a reasonable man would object to: To repeat, it is not unreasonably risky to send the boy to summer camp by airplane.

The same analysis applies to many cases of impossible attempts. The man who tried to kill his wife by witchcraft was engaging in conduct a reasonable man would not object to. It is not unreasonably risky to stick pins into someone's likeness. . . .

In the present case, did the defendant Omeira create an unreasonable risk that valuable artifacts of the native culture would be exported? Would a reasonable man have objected to his conduct? I believe so. The forgery was near-perfect. A reasonable man would certainly have been worried that what Omeira was trying to export was a real Decameron. Of course, if the forgery were terribly crude, so crude that any reasonable man could detect it, the answer would be different.

I conclude that the defendant's conviction should be affirmed.

[The court being evenly divided, Justice Middlebury will cast the deciding vote. I leave it to the reader to make Justice Middlebury's decision for him.]

REVIEW AND DISCUSSION QUESTIONS

1. Explain the basic facts of this case. Give examples of failed crimes that are crimes, and failed crimes that are not crimes.
2. What basic principle does Chief Justice Newson appeal to, and how does it apply to Jan Omeira? Why does Justice Henchard disagree with this principle?
3. Despite disagreeing with Newson's reasoning, Justice Henchard believes that Omeira should be acquitted. Explain and assess his position.
4. Why does Justice Farfrae criticize the distinction that Henchard draws between "legal impossibility" and "factual impossibility"? Is it a workable distinction? With whom do you agree about this—Henchard or Farfrae?
5. Farfrae says that one must ask two questions about a case like this. What are those questions, and how does Farfrae answer them?
6. Should Omeira be found guilty of attempting to export illegally "valuable artifacts of the native culture"? Explain your reasoning.

Drunkenness

People v. Koerber

As a general matter, voluntary intoxication is no defense to a crime. However, sometimes it can be relevant to whether the defendant had the specific intent necessary for a particular crime or degree of a crime. For example, as Judge Pound explains in his majority opinion, because murder in the first degree requires deliberation and premeditation, a jury might decide that because of his or her drunkenness the defendant lacked the necessary intent for that crime. In this case, defendant Koerber was charged under the felony murder rule, which holds a defendant strictly liable for first-degree murder if a death occurs in the course of a felony committed by the defendant (in this case, a robbery). In other words, if Koerber committed a robbery and someone died as a result of it, then Koerber is guilty of first-degree murder.

The trial judge instructed the jury that voluntary intoxication does not excuse a crime. This was an error, Pound argues, because Koerber's intoxication might be relevant—not directly to the murder charge—but to the jury's determining whether he had the specific intent to commit robbery, the underlying felony. If he lacked the intent necessary for that felony, the indictment for felony murder would fail. In their separate dissents, Justices Andrews and Crane argue that robbery requires only general criminal intent and that there is no evidence that Koerber was so intoxicated that he did not and could not form an intent to rob. Therefore, they contend, the trial court was right in refusing to submit the question of intoxication to the jury.

POUND, J.: Defendant was convicted of murder in the first degree for killing one Mahairas when engaged in the commission of a felony, to wit, the crime of robbery, upon the person killed. The homicide occurred about five o'clock in the morning of Sunday, March 21, 1926. Defendant testified in his own behalf in substance that he had been drinking heavily and that he was intoxicated at the time so that he did not know what he was doing. He narrated with considerable detail the occurrences of the night, both before and after the homicide, and the jury might properly have found that his intoxication was not such as to affect his responsibility for his acts had it been permitted to pass on the question. It was, however, instructed merely that voluntary intoxication would not excuse a criminal act committed while under the influence of intoxication. A request to charge the jury as to the different degrees of homicide and the rule of reasonable doubt as applied to lower degrees of crime was refused and also a request to instruct the jury as to intoxication as a defense substantially in the language of section 1220 of the Penal Law. The question of voluntary intoxication as defense for crime is thus fairly, although not with any great degree of precision, presented on the record, both on the facts and the law.

On the question of murder in the first degree where the killing is committed from a deliberate and premeditated design to effect the death of the person killed, it has been held that, when the case contains testimony which if believed might have led the jury to conclude that at the time of the shooting defendant was intoxicated to a greater or less degree, the question of intoxication should be presented and explained to the jury and, if the question is a serious one and such instructions are omitted, a judgment of conviction should be reversed even in the absence of a proper request by the counsel for the defendant.

The view is expressed that in felony murder the rule is different; that a refusal to instruct the jury as to the lower degrees of homicide is proper and that the jury should be instructed to find the defendant guilty of murder in the first degree or to acquit. Such is the rule only "where . . . no possible view of the facts would justify any other verdict except a conviction of the crime charged or an acquittal." We are,

244 N.Y. 147 (1926).

therefore, free to enter upon a consideration of voluntary intoxication as a fact for the consideration of the jury in such cases.

The effect of this section [Penal Law, section 1220] is discussed in *People v. Leonardi* (143 N.Y. 360) where the indictment was for a deliberate and premeditated killing. At common law, said the court, if a man made himself voluntarily drunk it was no excuse for any crime he might commit when he was in that condition, and he had to take the responsibility of his own voluntary act in becoming intoxicated. The fact was not to be considered even on the question of premeditation. Under the section, however, while no act of a person in a state of voluntary intoxication shall be deemed less criminal by reason of his having been in such condition, the judge should, whenever the actual existence of any particular purpose, motive or intent is a necessary element to constitute a particular species or degree of crime, allow the jury to take into consideration the question of intoxication in determining the existence of such particular intent and, if it should be of the opinion that the deliberation or premeditation necessary to constitute murder in the first degree did not exist, the crime might be reduced to a lower degree of murder or, in the absence of any intent to kill, then to manslaughter in some of its degrees. That is, the law does not, as in the case of insanity, deem the act less criminal by reason of the voluntary intoxication of the accused, but the jury may take into consideration the fact of intoxication in determining the element of the crime charged, not as calling for an acquittal, where the act charged, minus the particular intent, remains a criminal act, but as tending to reduce the character or grade of the offense in a proper case. . . .

In brief then, the rule may be thus stated: When criminal intent in general is all that need be established the drunken defendant is treated as if he knew the consequence of his acts; but where a particular or specific intent must be established, if the jury find that the mind of the defendant was so obscured by drink that he was incapable of forming that intent, it may justify itself in the reduction of a charge.

Generally speaking murder in the first or second degree connotes the specific or particular intent to kill, while manslaughter in the first or second degree is felonious homicide when the intent to kill is absent. But when one engaged in the commission of a felony, his mind being fatally bent on mischief but without a design to effect death, kills a human being, at common law, the killing is said to be with malice aforethought and so murder, and the Penal Law attaches to the act the consequences of murder in the first degree. The People on an indictment for felony murder may fail to establish that defendant was engaged in the commission of a felony, but may offer evidence tending to show that the homicide was committed by him when engaged in a misdemeanor, which would reduce the offense to manslaughter in the first degree, or in the commission of a trespass or other invasion of a private right, which would reduce it to manslaughter in the second degree. Where the facts would justify a verdict based on such a theory, the degrees of manslaughter should be submitted to the jury. Thus, where the felony charged in the indictment is burglary and the entry is under circumstances or in a manner not amounting to a burglary, the offense is unlawfully entering a building, which is a misdemeanor. Could it be said with reason that on an indictment for murder in the first degree for killing when engaged in the commission of burglary, the offense might not thus be reduced to manslaughter in the first degree? Or, if it were shown that the killing took place before a building was entered while defendant was engaged in a trespass or other invasion of a private right, that the offense might not be thus reduced to manslaughter in the second degree? . . .

We now come to consider the charge against defendant as affected by his intoxication. Disregarding for the moment the statutory modification of common-law rules and principles and the fine distinctions not applicable hereto, we find that the gist of robbery is larceny by force from the person, and that the gist of larceny is the taking and carrying away of personal property of another with the specific intent to steal such property. If on an indictment for larceny or robbery the testimony leaves uncertain the intent with which the accused took the property, the jury should be instructed to give the defendant the benefit of the doubt. That the intent is usually to be inferred from the act does not change the rule that a taking without intent to steal is not larceny at common law. No statutory larceny is involved in the consideration of the instant case.

Robbery, as thus defined, is "a particular species . . . of crime" of which "the actual existence of any particular . . . intent is a necessary element" within the meaning of Penal Law, section 1220. It follows that the jury should have been instructed "to take into consideration the fact that the accused was intoxicated at the time, in determining the . . . intent with which he committed the act." The fact that the accused did not have the intent to rob would not mean that an acquittal should follow. It is conceivable, for example, that a youth whose mind was befuddled by drink might intend merely to stage a hold-up and yet be guilty of some degree of homicide. He might be so frightened by resistance as to shoot in the heat of passion, the emotion of fright, and thus be guilty of manslaughter in the first degree, or conceivably, even of murder in the second degree. The jury would have to say under proper instructions as to the degrees of crime.

It is common knowledge that intoxicated men, although not in normal control of their faculties, do deliberate and premeditate and form a particular intent and commit criminal acts as they might not do if they were sober. Intoxication as such does not mitigate the offense. The question is not whether the accused was drunk but whether his intoxication was of such a character that it destroyed the power to form the particular intent which is a necessary element of the crime charged. The jury should, therefore, proceed with the utmost caution before arriving at the conclusion that the voluntary intoxication of an accused person has in any way altered the character or grade of his criminal act. His own evidence on the point need not be accepted as true even if uncontradicted. We may doubt whether the evidence of intoxication adduced by this defendant would carry sufficient weight with an intelligent jury to affect its verdict. It presented, however, a serious question affecting a substantial right which should not have been withheld from the consideration of triers of fact. We cannot say that, with proper instructions, "but one decision and that adverse to the defendant could reasonably have been reached." When the alternative presented was conviction of murder in the first degree or acquittal, a conscientious jury would scarcely bring itself to a verdict of not guilty in this case. If they had been instructed that other verdicts were permissible, they might or might not have found the defendant guilty of a lesser degree of felonious homicide. We, therefore, cannot overlook the failure of the court to give proper instructions . . . , as we might if we could reach the conclusion that there was a lack of sufficient evidence to go to the jury that defendant was, in the only relevant sense, too drunk to form the specific intent of committing robbery.

The judgment of conviction should be reversed and a new trial ordered.

ANDREWS, J. (dissenting): Under the charge of the court we are here concerned solely with robbery. If guilty of that crime the defendant was guilty of murder in the first degree. He could be guilty of nothing else. If not guilty of robbery the jury was told that they must acquit him. Surely he may not complain because in that case they were not permitted to convict him of some lesser degree of homicide.

To be guilty of robbery it is essential that the defendant should have acted with a general criminal intent. I assume that means the intent to take the property of the deceased in his presence, against his will, by means of force or fear. Ordinarily from the occurrence itself the jury might and should have inferred this intent. He may, however, negative this presumption. He may give evidence tending to show that he was insane at the time. Or that he was playing a joke on a friend. Or that he was so drunk that he could not and did not form any such intent. Thus a question of fact may be presented, and if the jury have a reasonable doubt caused by this explanation he should be acquitted.

The issue is as to the existence of this intent, and the testimony offered by the defendant must be relevant to this precise issue. If it does not tend to show that the accused failed to know the nature and quality of his act or to show that he did not know that it was wrong, there is no question of insanity to be submitted to the jury. If it does not tend to show that he was so intoxicated that he did not and could not form an intent to rob, again there is no question for them.

In my judgment testimony of intoxication to the extent that would make the defendant's intent a question of fact is wholly absent from this case. It should, at least, tend to show at the time of the

robbery a complete absence of conscious volition—an ignorance of his acts and purposes. Therefore, the trial court was right in refusing to submit any such question to the jury. The general rule that "no act committed by a person while in a state of voluntary intoxication, shall be deemed less criminal by reason of his having been in such a condition" is applicable.

It is quite true that in a limited number of cases another rule applies. Now and then crimes are divided into degrees or species dependent upon the motive, purpose, or intent of the criminal. His legal as well as his moral guilt varies with his mental attitude. Such a division is made as regards some kinds of homicide. So as to assault. Ordinarily, however, while other crimes may be divided into degrees, these degrees are not made dependent on any such purpose, motive or intent. There are degrees in arson, in burglary, in larceny, in forgery, in robbery itself; but no such test of guilt applies. The degree depends upon other circumstances. To convict there must be merely the usual general intent to commit the crime.

It is only in the exceptional cases of the first class that the jury may take into consideration the fact that the accused was to some extent intoxicated at the time in determining the question of the degree of guilt.

In my opinion the judgment of conviction should be affirmed.

CRANE, J. (dissenting): I concur with JUDGE ANDREWS, but desire to add the following:

That the defendant shot and killed the deceased has been proved beyond any reasonable doubt. The evidence shows that the killing happened while the defendant was committing the crime of robbery—it was a holdup case—in which the defendant at the point of a pistol attempted to rob the deceased. The attempted defense was that the deceased attacked the defendant leaving the store, and in the mix-up the defendant's pistol went off accidentally. No one takes any stock in this defense, not even the defendant's counsel, who admitted on the argument that the defendant was guilty of a crime.

A discussion has arisen over the effect of liquor upon the defendant and whether the judge fully covered the law regarding intoxication as affecting crime and the degrees of crime. In my judgment the discussion is immaterial to this case, as the evidence clearly shows that the defendant, if he had been drinking, was not so intoxicated as to lessen the crime, but rather to aggravate the seriousness of it. In the first place, he is the only one who states that he had been drinking, and then on the stand as a witness remembers every detail of his conduct and actions on the night in question. He boldly states how a few hours previous to killing the deceased, he had held up a restaurant keeper at the point of a pistol—pointed a gun at him and told him to hold up his hands. The money he took he says he did not intend to steal, that it was a kind of joke—the kind that the people of this state do not appreciate. It was no joke to the restaurant keeper. Later, the defendant details how he entered the place of the deceased and what happened there. Surely the defendant, when he had sufficient mind and intelligence to recollect months afterwards all that happened on the night of the shooting, was not under such influence from liquor as to lessen in any way his criminal acts, or deprive him of criminal intent. His intoxication may have excited him and weakened his judgment, but it did not absolve him from the responsibility which the law attaches to one who kills another in an attempt to take his property.

I am for affirmance for these reasons.

REVIEW AND DISCUSSION QUESTIONS

1. Explain the facts of this case. How does the felony murder rule apply to it?
2. Cardozo states that in New York the common-law rule that voluntary intoxication is no excuse for a crime has been modified by statute. Explain. What is the rule now?
3. Cardozo writes that "intoxication as such does not mitigate the offense" but it can "destroy the power to form the particular intent" necessary to the crime. Explain.
4. According to Cardozo, what is the "gist of robbery" and what intent does it require? Could someone who performs the physical act of robbing someone be so drunk as to lack the necessary *mens rea* for robbery?

5. According to Andrews, whereas some crimes (like homicide) are divided into degrees based on the defendant's specific mental attitude, in most crimes only general criminal intent is required and the degrees of the crime depend on other factors. Explain.
6. Andrew and Crane believe that there is no serious evidence that Koerber was so intoxicated as to have been unable to commit robbery. Should this have been a factual question left for the jury to determine, or was the trial judge right not to have potentially confused the jury with this issue?
7. Should Koerber have been given a new trial, as Cardozo ordered?
8. According to the felony murder rule, if someone is killed accidentally in the course of a felony, then (1) what might have been manslaughter is increased to murder and (2) all the participants in the felony are held responsible for the murder. What can be said for and against the felony murder rule? Explain why you do or do not support it.

The Battered Woman Syndrome

State v. Leidholm

Among the most controversial murder cases are those involving female defendants suffering from so-called battered woman syndrome. In this North Dakota case, Janice Leidholm appealed her conviction for killing her abusive husband. In reviewing the case, the Supreme Court of North Dakota discusses the battered woman syndrome and its relationship to the law of self-defense and "objective" versus "subjective" standards of reasonableness.

VANDE WALLE, JUSTICE: Janice Leidholm was charged with murder for the stabbing death of her husband, Chester Leidholm, in the early morning hours of August 7, 1981, at their farm home near Washburn. She was found guilty by a McLean County jury of manslaughter and was sentenced to five years' imprisonment in the State Penitentiary with three years of the sentence suspended. Leidholm appealed from the judgment of conviction. We reverse and remand the case for a new trial.

I

According to the testimony, the Leidholm marriage relationship in the end was an unhappy one, filled with a mixture of alcohol abuse, moments of kindness toward one another, and moments of violence. The alcohol abuse and violence was exhibited by both parties on the night of Chester's death.

Early in the evening of August 6, 1981, Chester and Janice attended a gun club party in the city of Washburn where they both consumed a large amount of alcohol. On the return trip to the farm, an argument developed between Janice and Chester which continued after their arrival home just after midnight. Once inside the home, the arguing did not stop; Chester was shouting, and Janice was crying.

At one point in the fighting, Janice tried to telephone Dave Vollan, a deputy sheriff of McLean County, but Chester prevented her from using the phone by shoving her away and pushing her down. At another point, the argument moved outside the house, and Chester once again was pushing Janice to the ground. Each time Janice attempted to get up, Chester would push her back again.

A short time later, Janice and Chester re-entered their home and went to bed. When Chester fell asleep, Janice got out of bed, went to the kitchen,

334 N.W. 2d. 811 (1983).

and got a butcher knife. She then went back into the bedroom and stabbed Chester. In a matter of minutes Chester died from shock and loss of blood.

II

. . . The first, and controlling, issue we consider is whether or not the trial court correctly instructed the jury on self-defense.

A defense of justification is the product of society's determination that the *actual existence* of certain circumstances will operate to make proper and legal what otherwise would be criminal conduct. A defense of excuse, contrarily, does not make legal and proper conduct which ordinarily would result in criminal liability; instead, it openly recognizes the criminality of the conduct but excuses it because the actor believed that circumstances actually existed which would justify his conduct when in fact they did not. In short, had the facts been as he supposed them to be, the actor's conduct would have been justified rather than excused. . . .

In the context of self-defense, this means that a person who believes that the force he uses is necessary to prevent imminent unlawful harm is *justified* in using such force if his belief is a *correct* belief; that is to say, if his belief corresponds with what actually is the case. If, on the other hand, a person *reasonably* but incorrectly believes that the force he uses is necessary to protect himself against imminent harm, his use of force is *excused.*

Courts have traditionally distinguished between standards of reasonableness by characterizing them as either "objective" or "subjective." An objective standard of reasonableness requires the factfinder to view the circumstances surrounding the accused at the time he used force from the standpoint of a hypothetical reasonable and prudent person. Ordinarily, under such a view, the unique physical and psychological characteristics of the accused are not taken into consideration in judging the reasonableness of the accused's belief.

This is not the case, however, where a subjective standard of reasonableness is employed. See *State v. Wanrow,* 88 Wash. 2d 221, 559 P.2d 548 (1977). Under the subjective standard the issue is not whether the circumstances attending the accused's use of force would be sufficient to create in the mind of a reasonable and prudent person the belief that the use of force is necessary to protect himself against immediate unlawful harm, but rather whether the circumstances are sufficient to induce in the accused an honest and reasonable belief that he must use force to defend himself against imminent harm.

Because (1) the law of self-defense as developed in past decisions of this court has been interpreted to require the use of a subjective standard of reasonableness, and (2) we agree with the court in *Hazlett* that a subjective standard is the more just, and (3) our current law of self-defense . . . does not require a contrary conclusion, that is to say, our current law of self-defense is consistent with either a subjective or objective standard, we now decide that the finder of fact must view the circumstances attending an accused's use of force from the standpoint of the accused to determine if they are sufficient to create in the accused's mind an honest and reasonable belief that the use of force is necessary to protect himself from imminent harm. . . .

The practical and logical consequence of this interpretation is that an accused's actions are to be viewed from the standpoint of a person whose mental and physical characteristics are like the accused's and who sees what the accused sees and knows what the accused knows. For example, if the accused is a timid, diminutive male, the factfinder must consider these characteristics in assessing the reasonableness of his belief. If, on the other hand, the accused is a strong, courageous, and capable female, the factfinder must consider these characteristics in judging the reasonableness of her belief.

In its statement of the law of self-defense, the trial court instructed the jury:

> The circumstances under which she acted must have been such as to produce in the mind of reasonably prudent persons, regardless of their sex, similarly situated, the reasonable belief that the other person was then about to kill her or do serious bodily harm to her.

In view of our decision today, the court's instruction was a misstatement of the law of self-defense. A correct statement of the law to be applied in a case of self-defense is:

> [A] defendant's conduct is not to be judged by what a reasonably cautious person might or might not do or consider necessary to do under the like circumstances,

but what he himself in good faith honestly believed and had reasonable ground to believe was necessary for him to do to protect himself from apprehended death or great bodily injury.

The significance of the difference in viewing circumstances from the standpoint of the "defendant alone" rather than from the standpoint of a "reasonably cautious person" is that the jury's consideration of the unique physical and psychological characteristics of an accused allows the jury to judge the reasonableness of the accused's actions against the accused's subjective impressions of the need to use force rather than against those impressions which a jury determines that a hypothetical reasonably cautious person would have under similar circumstances. . . .

Hence, a correct statement of the law of self-defense is one in which the court directs the jury to assume the physical and psychological properties peculiar to the accused, viz., to place itself as best it can in the shoes of the accused, and then decide whether or not the particular circumstances surrounding the accused at the time he used force were sufficient to create in his mind a sincere and reasonable belief that the use of force was necessary to protect himself from imminent and unlawful harm. . . .

Leidholm argued strongly at trial that her stabbing of Chester was done in self-defense and in reaction to the severe mistreatment she received from him over the years. Because the court's instruction in question is an improper statement of the law concerning a vital issue in Leidholm's defense, we conclude it amounts to reversible error requiring a new trial.

The expert witness in this case testified that Janice Leidholm was the victim in a battering relationship which caused her to suffer battered woman syndrome manifested by (1) a psychological condition of low self-esteem and (2) a psychological state of "learned helplessness." . . .

The instruction on battered woman syndrome was designed to support Leidholm's claim of self-defense by focusing the jury's attention on the psychological characteristics common to women who are victims in abusive relationships, and by directing the jury that it may consider evidence that the accused suffered from battered woman syndrome in determining whether or not she acted in self-defense. The instruction correctly points out that battered woman syndrome is not of itself defense. In other words, "The existence of the syndrome in a marriage does not of itself establish the legal right of the wife to kill the husband; the evidence must still be considered in the context of self-defense." . . .

There is nothing in the proposed instruction at issue which would add to or significantly alter a correct instruction on the law of self-defense. The jury's use of a subjective standard of reasonableness in applying the principles of self-defense to the facts of a particular case requires it to consider expert testimony, once received in evidence, describing battered woman syndrome and the psychological effects it produces in the battered spouse when deciding the issue of the existence and reasonableness of the accused's belief that force was necessary to protect herself from imminent harm. If an instruction given is modeled after the law of self-defense which we adopt today, the court need not include a specific instruction on battered woman syndrome in its charge to the jury. . . .

The judgment of conviction is reversed and the case is remanded . . . for a new trial.

REVIEW AND DISCUSSION QUESTIONS

1. Under what circumstances does self-defense provide a justification, and under what circumstances does it provide an excuse, to a homicide charge? In general, what's the difference between saying that a defendant's action was justified and saying that it should be excused?
2. In your view, was Janice Leidholm's slaying of her husband justified? Explain why or why not. If not, should it be excused?
3. What's the difference between an "objective" and a "subjective" standard of reasonableness? How does the difference affect cases involving defendants suffering from battered woman syndrome?

4. The court believes that the use of force can be reasonable even though a (hypothetical) reasonable person would not have used force in those circumstances. Do you agree?
5. Compare the trial court's statement of the law of self-defense with that provided by the court in this case. Is the stance of the North Dakota Supreme Court too permissive? Is it more just to defendants? To victims? To society?
6. According to the standard employed by the court in this case, a defendant can plead self-defense as an excuse if his or her belief was honest and in good faith and if it was "reasonable." But suppose the defendant's belief was honest and in good faith, but was unreasonable, even when judged from his or her own standpoint. Should the person be excused? Explain why or why not.

Rape, Consent, and Mens Rea

Regina v. Morgan

A finding of "mens rea" or guilty mind is almost always necessary before a person can be convicted of a crime. Killing another in an automobile accident, for instance, is not the same as committing a murder because the driver of the car did not intend to kill, or even foresee the death. In this much discussed case, the question before the House of Lords is whether someone who unreasonably yet genuinely believed a woman was consenting can be convicted of rape, or whether they are excused on the ground that they lacked mens rea.

The defendant Morgan and three other defendants were convicted of the forcible rape of Morgan's wife. Morgan's liability rested on his having aided and abetted the three others. The Court of Appeals affirmed all the convictions but certified the following question to the House of Lords:

> Whether in rape the defendant can properly be convicted notwithstanding that he in fact believed that the woman consented, if such a belief was not based on reasonable grounds.

[The applicable statute stated simply: "*Rape*—(1) It is a felony for a man to rape a woman."]

Lord Hailsham of St. Marylebone: The question arises in the following way. The appellant Morgan and his three co-appellants, who were all members of the RAF, spent the evening of 15th August 1973 in one another's company. The appellant Morgan was significantly older than the other three, and considerably senior to them in rank. He was . . . married to the alleged victim, but not, it seems, at the time habitually sleeping in the same bed. [B]y the time the appellants arrived at Morgan's house, Mrs. Morgan was already in bed and asleep, until she was awoken by their presence.

According to the version of the facts which she gave in evidence, and which was evidently accepted by the jury, she was aroused from her sleep, . . . held by each of her limbs, . . . while each of the three young appellants in turn had intercourse with her in the presence of the others. . . .

According to Mrs. Morgan she consented to none of this and made her opposition to what was being done very plain indeed. . . .

All four appellants explained in the witness box that they had spent the evening together in Wolverhampton, and by the time of the alleged offences had had a good deal to drink. Their original intention had been to find some women in the town but, when this failed, Morgan made the surprising

A.C. 182 (1976).

suggestion to the others that they should all return to his home and have sexual intercourse with his wife. According to the three younger appellants (but not according to Morgan, who described this part of their story as "lying") Morgan told them that they must not be surprised if his wife struggled a bit, since she was "kinky" and this was the only way in which she could get "turned on." However this may be, it is clear that Morgan did invite his three companions home in order that they might have sexual intercourse with his wife and, no doubt, he may well have led them in one way or another to believe that she would consent to their doing so. This, however, would only be matter predisposing them to believe that Mrs. Morgan consented, and would not in any way establish that, at the time, they believed she did consent whilst they were having intercourse.

I needed not enter into the details of what the appellants said happened after they had arrived at the house. As I have said they admitted that some degree of struggle took place in the wife's bedroom. But all asserted that after she got into the double bedroom she not merely consented to but actively co-operated with and enjoyed what was being done. . . .

The choice before the jury was thus between two stories each wholly incompatible with the other, and in my opinion it would have been quite sufficient for the judge, after suitable warnings about the burden of proof, corroboration, separate verdicts and the admissibility of the statements only against the makers, to tell the jury that they must really choose between the two versions. . . .

The certified question arises because counsel for the appellants raised the question whether, even if the victim [objected], the appellants may . . . have honestly believed that she did [not]. [I]n the summing-up [to the jury] . . . [t]he learned judge said:

> [T]he prosecution have to prove that each defendant intended to have sexual intercourse with this woman without her consent. Not merely that he intended to have intercourse with her but that he intended to have intercourse without her consent. Therefore if the defendant believed or may have believed that Mrs. Morgan consented to him having sexual intercourse with her, then there would be no such intent in his mind and he would be not guilty of the offence of rape, but such a belief must be honestly held by the defendant in the first place. He must really believe that. And, secondly, his belief must be a reasonable belief; such a belief as a reasonable man would entertain if he applied his mind and thought about the matter. It is not enough for a defendant to rely upon a belief, even though he honestly held it, if it was completely fanciful; contrary to every indication which could be given which would carry some weight with a reasonable man.

It is on the second proposition about the mental element that the appellants concentrate their criticism. An honest belief in consent, they contend, is enough. It matters not whether it be also reasonable. . . .

The beginning of wisdom in all the "mens rea" cases to which our attention was called is, as was pointed out by Stephen, J., in *R. v. Tolson,* [23 Q.B.D. 168 (1889)], that "mens rea" means a number of quite different things in relation to different crimes. . . .

Once one has accepted, what seems to me abundantly clear, that the prohibited act in rape is non-consensual sexual intercourse, and that the guilty state of mind is an intention to commit it, it seems to me to follow as a matter of inexorable logic that there is no room either for a "defence" of honest belief or mistake, or of a defence of honest and reasonable belief and mistake. Either the prosecution proves that the accused had the requisite intent, or it does not. In the former case it succeeds, and in the latter it fails. Since honest belief clearly negates intent, the reasonableness or otherwise of that belief can only be evidence for or against the view that the belief and therefore the intent was actually held. . . .

For the above reasons I would answer the question certified in the negative, but would apply the proviso to s. 2(1) of the Criminal Appeal Act 1968 on the ground that no miscarriage of justice has or conceivably could have occurred. In my view, therefore these appeals should be dismissed.

Lord Fraser of Tulleybelton: The argument for the Crown in support of an affirmative answer to the question in this case was not supported by any English decision on rape. It was supported by reference to English decisions in relation to other offences which are more or less analogous to rape.

. . . The English case upon which most reliance was placed was *Reg. v. Tolson,* 23 Q.B.D. 168, which

was concerned with bigamy, and which decided that a bona fide belief *on reasonable grounds* in the death of the husband at the time of the second marriage afforded a good defence to the indictment for bigamy. The main argument in the case was concerned with the question whether a mistaken belief could be a defence to a charge of bigamy at all, and comparatively little attention was given to the subsidiary point of whether the belief had to be based upon reasonable grounds. The case seems to me therefore of only limited assistance for the present purpose. . . .

The difficulty of arguing by analogy from one offence to another is strikingly illustrated by reference to the case of *Reg. v. Prince* (1875) 13 Cox C.C. 138. That case dealt with abduction of a girl under the age of 16, an offence created by section 55 of the Act of 1861. Bramwell, B., with whom five other judges concurred, held that a mistaken and reasonable belief by the defendant that the abducted girl was aged 16 or more was no excuse, because abduction of a young girl was immoral as well as illegal, although a mistaken and reasonable belief by the defendant that he had the consent of the girl's father would have been an excuse. If such differences can exist about mistaken beliefs of different facts in one offence, it is surely dangerous to argue from one offence to another. No doubt a rapist, who mistakenly believes that the woman is consenting to intercourse, must be behaving immorally, by committing fornication or adultery. But those forms of immoral conduct are not intended to be struck at by the law against rape; indeed, they are not now considered appropriate to be visited with penalties of the criminal law at all. There seems therefore to be no reason why they should affect the consequences of the mistaken belief. . . .

For these reasons, I am of the opinion that there is no authority which compels me to answer the question in what I would regard as an illogical way. I would therefore answer the question in the negative—i.e., in favour of the accused. But for the reasons stated by . . . Lord Hailsham . . . I would apply the proviso to the Criminal Appeal Act of 1968 . . . and I would refuse the appeal.

Appeals Dismissed.

REVIEW AND DISCUSSION QUESTIONS

1. What are the facts in this case? What were the defendants accused of doing?
2. At the trial, the judge instructed the jurors that they had to make two determinations to convict. What were the two? What was the legal question before the House of Lords?
3. How is intent relevant to a criminal charge of rape? Lord Halisham argues that it is irrelevant whether the defendants' belief in this case was reasonable or unreasonable. Explain why. Do you agree or disagree with his reasoning?
4. On appeal the prosecution rested its argument in part on precedents established in other crimes somewhat analogous to rape. Lord Fraser states that he is not persuaded. Explain why.
5. From a moral point of view, is it correct to say that the defendants did nothing wrong? Why do you think the Lords said that the conviction should stand?

Rape, Force, and Consent

Susan Estrich

The criminal law's handling of rape has long been rife with sexism, with the law appearing to be more concerned to protect men from unfair convictions than to do justice to rape victims. In this essay, Susan Estrich, professor of law at the University of Southern California, explores the ways in which male standards and traditional conceptions of sex roles have shaped the crime of rape. Unlike their

British counterparts, American courts have tended to omit *mens rea* from the definition of rape—thus treating what the man knew or should have known about the victim's consent as irrelevant. But the results of this have not been favorable to women. Instead, by making force or threat of force central to rape and by requiring that the victim prove lack of consent by her resistance, the courts have made the burden of proving rape unacceptably high.

THE DEFINITION OF RAPE

The traditional way of defining a crime is by describing the prohibited act (*actus reus*) committed by the defendant and the prohibited mental state (*mens rea*) with which he must have done it. We ask: What did the defendant do? What did he know or intend when he did it?

The definition of rape stands in striking contrast to this tradition, because courts, in defining the crime, have focused almost incidentally on the defendant—and almost entirely on the victim. It has often been noted that, traditionally at least, the rules associated with the proof of a rape charge—the corroboration requirement, the requirement of cautionary instructions, and the fresh complaint rule—as well as the evidentiary rules relating to prior sexual conduct by the victim, placed the victim as much on trial as the defendant. Such a reversal also occurs in the course of defining the elements of the crime. *Mens rea,* where it might matter, is all but eliminated; prohibited force tends to be defined according to the response of the victim; and nonconsent—the *sine qua non* of the offense—turns entirely on the victim's response.

But while the focus is on the female victim, the judgment of her actions is entirely male. If the issue were what the defendant knew, thought, or intended as to key elements of the offense, this perspective might be understandable; yet the issue has instead been the appropriateness of the woman's behavior, according to male standards of appropriate female behavior.

To some extent, this evaluation is but a modern response to the longstanding suspicion of rape victims. As Matthew Hale put it three centuries ago: "Rape is . . . an accusation easily to be made and hard to be proved, and harder to be defended by the party accused, tho never so innocent."[1]

But the problem is more fundamental than that. Apart from the woman's conduct, the law provides no clear, working definition of rape. This rather conspicuous gap in the law of rape presents substantial questions of fair warning for men, which the law not so handily resolves by imposing the burden of warning them on women.

At its simplest, the dilemma lies in this: If nonconsent is essential to rape (and no amount of force or physical struggle is inherently inconsistent with lawful sex), and if no sometimes means yes, and if men are supposed to be aggressive in any event, how is a man to know when he has crossed the line? And how are we to avoid unjust convictions?

This dilemma is hardly inevitable. Partly, it is a product of the way society (or at least a powerful part of it) views sex. Partly, it is a product of the lengths to which the law has gone to enforce and legitimize those views. We could prohibit the use of force and threats and coercion in sex, regardless of "consent." We could define consent in a way that respected the autonomy of women. Having chosen neither course, however, we have created a problem of fair warning, and force and consent have been defined in an effort to resolve this problem.

Usually, any discussion of rape begins (and ends) with consent. I begin instead with *mens rea,* because if unjust punishment of the blameless man is our fear (as it was Hale's), then *mens rea* would seem an appropriate place to start addressing it. At least a requirement of *mens rea* would avoid unjust convictions without adjudicating the "guilt" of the victim. It could also be the first step in expanding liability beyond the most traditional rape. . . .

MENS REA

It is difficult to imagine any man engaging in intercourse accidentally or mistakenly. It is just as difficult to imagine an accidental or mistaken use of force, at least as force is conventionally defined. But it is not at all difficult to imagine cases in which a man might claim that he did not realize that the woman was not consenting to sex. He may have been mistaken in

assuming that no meant yes. He may not have bothered to inquire. He may have ignored signs that would have told him that the woman did not welcome his forceful penetration.

In doctrinal terms, such a man could argue that his mistake of fact should exculpate him because he lacked the requisite intent or *mens rea* as to the woman's required nonconsent.[2] American courts have altogether eschewed the *mens rea* or mistake inquiry as to consent, opting instead for a definition of the crime of rape that is so limited that it leaves little room for men to be mistaken, reasonably or unreasonably, as to consent. The House of Lords, by contrast, has confronted the question explicitly and, in its leading case, has formally restricted the crime of rape to men who act recklessly, a state of mind defined to allow even the unreasonably mistaken man to avoid conviction.

This Section argues that the American courts' refusal to confront the *mens rea* problem works to the detriment of the victim. In order to protect men from unfair convictions, American courts end up defining rape with undue restrictiveness. . . .

In defining the crime of rape, most American courts have omitted *mens rea* altogether. In Maine, for example, the Supreme Judicial Court has held that there is no *mens rea* requirement at all for rape. In Pennsylvania, the Superior Court held in 1982 that even a reasonable belief as to the victim's consent would not exculpate a defendant charged with rape. In 1982 the Supreme Judicial Court of Massachusetts left open the question whether it would recognize a defense of reasonable mistake of fact as to consent, but it rejected the defendant's suggestion that any mistake, reasonable or unreasonable, would be sufficient to negate the required intent to rape; such a claim was treated by the court as bordering on the ridiculous. The following year the court went on to hold that a specific intent that intercourse be without consent was not an element of the crime of rape; that decision has since been construed to mean that there is no intent requirement at all as to consent in rape cases.

To treat what the defendant intended or knew or even should have known about the victim's consent as irrelevant to his liability sounds like a result favorable to both prosecution and women as victims. But experience makes all too clear that it is not. To refuse to inquire into *mens rea* leaves two possibilities: turning rape into a strict liability offense where, in the absence of consent, the man is guilty of rape regardless of whether he (or anyone) would have recognized nonconsent in the circumstances; or defining the crime of rape in a fashion that is so limited that it would be virtually impossible for any man to be convicted where he was truly unaware or mistaken as to nonconsent. In fact, it is the latter approach which has characterized all of the older, and many of the newer, American cases. In practice, abandoning *mens rea* produces the worst of all possible worlds: The trial emerges not as an inquiry into the guilt of the defendant (is he a rapist?) but of the victim (was she really raped? did she consent?). The perspective that governs is therefore not that of the woman, nor even of the particular man, but of a judicial system intent upon protecting against unjust conviction, regardless of the dangers of injustice to the woman in the particular case.

The requirement that sexual intercourse be accompanied by force or threat of force to constitute rape provides a man with some protection against mistakes as to consent. A man who uses a gun or knife against his victim is not likely to be in serious doubt as to her lack of consent, and the more narrowly force is defined, the more implausible the claim that he was unaware of nonconsent.

But the law's protection of men is not limited to a requirement of force. Rather than inquire whether the man believed (reasonably or unreasonably) that his victim was consenting, the courts have demanded that the victim demonstrate her nonconsent by engaging in resistance that will leave no doubt as to nonconsent. The definition of nonconsent as resistance—in the older cases, as utmost resistance, while in some more recent ones, as "reasonable" physical resistance—functions as a substitute for *mens rea* to ensure that the man has notice of the woman's nonconsent.

The choice between focusing on the man's intent or focusing on the woman's is not simply a doctrinal flip of the coin.

First, the inquiry into the victim's nonconsent puts the woman, not the man, on trial. Her intent, not his, is disputed; and because her state of mind is key, her sexual history may be considered relevant (even though utterly unknown to the man). Considering consent from *his* perspective, by

contrast, substantially undermines the relevance of the woman's sexual history where it was unknown to the man.

Second, the issue for determination shifts from whether the man is a rapist to whether the woman was raped. A verdict of acquittal thus does more than signal that the prosecution has failed to prove the defendant guilty beyond a reasonable doubt; it signals that the prosecution has failed to prove the woman's sexual violation—her innocence—beyond a reasonable doubt. Thus, as one dissenter put it in disagreeing with the affirmance of a conviction of rape: "The majority today . . . declares the innocence of an at best distraught young woman."[3] Presumably, the dissenter thought the young woman guilty.

Third, the resistance requirement is not only ill conceived as a definition of nonconsent, but is an overboard substitute for *mens rea* in any event. Both the resistance requirement and the *mens rea* requirement can be used to enforce a male perspective on the crime, but while *mens rea* might be justified as protecting the individual defendant who has not made a blameworthy choice, the resistance standard requires women to risk injury to themselves in cases where there may be no doubt as to the man's intent or blameworthiness. The application of the resistance requirement has not been limited to cases in which there was uncertainty as to what the man thought, knew or intended; it has been fully applied in cases where there can be no question that the man knew that intercourse was without consent. Indeed, most of the cases that have dismissed claims that *mens rea* ought to be required have been cases where both force and resistance were present, and where there was no danger of any unfairness.

Finally, by ignoring *mens rea,* American courts and legislators have imposed limits on the fair expansion of our understanding of rape. As long as the law holds that *mens rea* is not required, and that no instructions on intent need by given, pressure will exist to retain some form of resistance requirement and to insist on force as conventionally defined in order to protect men against conviction for "sex." Using resistance as a substitute for *mens rea* unnecessarily and unfairly immunizes those men whose victims are afraid enough, or intimidated enough, or, frankly, smart enough, not to take the risk of resisting physically. In doing so, the resistance test may declare the blameworthy man innocent and the raped woman guilty. . . .

In situations where a "reasonable man" would have known that the woman was not consenting, most defendants will face great difficulty in arguing that they were honestly mistaken or inadvertent as to consent. . . .

My view is that such a "negligent rapist" should be punished, albeit—as in murder—less severely than the man who acts with purpose or knowledge, or even knowledge of the risk. First, he is sufficiently blameworthy for it to be just to punish him. Second, the injury he inflicts is sufficiently grave to deserve the law's prohibition. . . .

In holding a man to such a standard of reasonableness, the law signifies that it considers a woman's consent to sex to be significant enough to merit a man's reasoned attention. In effect, the law imposes a duty on men to open their eyes and use their heads before engaging in sex—not to read a woman's mind, but to give her credit for knowing her own mind when she speaks it. The man who has the inherent capacity to act reasonably, but fails to do so, has made the blameworthy choice to violate this duty. . . .

FORCE AND THREATS

This Section examines two views of force in human relations. The first understands force as most schoolboys do on the playground: Force is when he hits me; resistance is when I hit back. That is the definition of force traditionally enforced in rape cases. A second understanding of force, not acknowledged in the law of rape, recognizes that bodily integrity means more than freedom from the force of fists, that power can be exercised without violence, and that coercion is not limited to what boys do in schoolyards.

Virtually every jurisdiction has traditionally made "force" or "threat of force" an element of the crime of rape. Where a defendant threatens his victim with a deadly weapon, beats her, or threatens to hurt her, and then proceeds immediately to have sex, few courts have difficulty finding that force is present. These facts fit the schoolboy definition of

force. But when some time elapses between the force and intercourse, when the force is more of the variety considered "incidental" to sex, or when the situation is threatening but no explicit threat of harm is communicated, "force" as defined and required by the criminal law may not be present at all. In such cases, the law fails to recognize, let alone protect, a woman's interest in bodily integrity. . . .

In *State v. Alston,*[4] Mr. Alston and the victim had been involved in a "consensual" relationship for six months. That relationship admittedly involved "some violence" by the defendant and some passivity by the victim. The defendant would strike the victim when she refused to give him money or refused to do what he wanted. As for sex, the court noted that "she often had sex with the defendant just to accommodate him. On those occasions, she would stand still and remain entirely passive while the defendant undressed her and had intercourse with her." This was their "consensual" relationship. It ended when, after being struck by the defendant, the victim left him and moved in with her mother.

A month later, the defendant came to the school which the victim attended, blocked her path, demanded to know where she was living and, when she refused to tell him, grabbed her arm and stated that she was coming with him. The victim told the defendant she would walk with him if he released her arm. They then walked around the school and talked about their relationship. At one point, the defendant told the victim he was going to "fix" her face; when told that their relationship was over, the defendant stated that he had a "right" to have sex with her again. The two went to the house of a friend. The defendant asked her if she was "ready," and the victim told him she did not want to have sexual relations. The defendant pulled her up from the chair, undressed her, pushed her legs apart, and penetrated her. She cried.

The defendant was convicted of rape, and his conviction was affirmed by the intermediate court of appeals. On appeal, the North Carolina Supreme Court agreed that the victim was not required to resist physically to establish nonconsent: The victim's testimony that she did not consent was "unequivocal" and her testimony provided substantial evidence that the act of sexual intercourse was against her will.

But the North Carolina Supreme Court nonetheless reversed on the ground that, even viewing the evidence in the light most favorable to the state, the element of force had not been established by substantial evidence. The victim did not "resist"—physically, at least. And her failure to resist, in the court's evaluation, was not a result of what the defendant did before penetration. Therefore, there was no "force."

The force used outside the school, and the threats made on the walk, "although they may have induced fear," were considered to be "unrelated to the act of sexual intercourse." Indeed, the court emphasized that the victim testified that it was not what the defendant said that day, but her experience with him in the past, that made her afraid. Such past experience was deemed irrelevant. . . .

Later in 1984, the North Carolina Court of Appeals applied *Alston* to another case where the defendant and the victim knew each other and had had previous sexual relations. In this case, however, the parties were not "boyfriend" and "girlfriend." They were a father and his 15-year-old daughter.

The defendant in *State v. Lester*[5] was the father of three daughters and a son. Prior to the parents' divorce, the defendant frequently beat the children's mother in their presence. He also beat his girlfriend and his son. He had a gun and on one occasion pointed it at his children. He engaged in sexual activity with all three of his daughters. He first had sexual relations with the daughter whose rape was at issue when she was 11 years old. Her mother found out and confronted the defendant. He swore never to touch her again, and then threatened to kill both mother and daughter if they told anyone of his actions. On both of the occasions in question, the victim initially refused her father's demand to take her clothes off and "do it." In both cases, she complied when the demand was repeated and she sensed that her father was becoming angry. The court held that the defendant could be convicted of incest, but not of rape.

. . . In each case [*Lester* and *Alston*], the court says—and this is explicit, not implicit—that sex was without the woman's consent. It also says that there was no force. In other words, the woman was not forced to engage in sex, but the sex she engaged in was against her will.

Such a paradox is almost inevitable if one adopts, and then enforces, the most traditional male notion of a fight as the working definition of "force." In a fight, you hit your assailant with your fists or your elbows or your knees. In a fight, the one attacked fights back. In these terms, there was no fight in *Alston.* Therefore, there was no force.

I am not at all sure how the judges who decided *Alston* would explain the victim's simultaneous refusal to consent and failure to resist. For myself, it is not at all difficult to understand that a woman who had been repeatedly beaten, who had been a passive victim of both violence and sex during the "consensual" relationship, who had sought to escape from the man, who is confronted and threatened by him, who summons the courage to tell him their relationship is over only to be answered by his assertion of a "right" to sex—a woman in such a position would not fight. She wouldn't fight; she might cry. Hers is the reaction of "sissies" in playground fights. Hers is the reaction of people who have already been beaten, or who never had the power to fight in the first instance. Hers is, from my reading, the most common reaction of women to rape. It certainly was mine. . . .

Unable to understand force as the power one need not use (at least physically), courts are left . . . to look for threats of force. Technically, these threats of force may be implicit as well as explicit. But implicit to whom? That a woman feels genuinely afraid, that a man has created the situation that she finds frightening, even that he had done it intentionally in order to secure sexual satisfaction, may not be enough to constitute the necessary force or even implicit threat of force which earns bodily integrity any protection under the law of rape. . . .

The simplest approach would be to ask whether this woman's will to resist was in fact overcome by this defendant's actions. Is she lying, or did she submit because she was truly frightened? If she is not lying—and none of the courts suggested that any of the women in these cases were actually lying—then affirm the conviction. But what about the poor man who didn't realize that the woman was overcome by fear of him, rather than desire for him? Properly regarded, such a man lacks *mens rea* as to force or consent.

A second approach resolves that problem without relying explicitly on *mens rea.* It asks instead: Were the defendant's acts and behavior intended to overcome this woman's will to resist? Under such a standard, . . . Mr. Lester [and] Mr. Alston . . . will have a hard time claiming that they didn't mean to succeed, and that success was not defined as creating a situation that would frighten the woman into submission.

A third approach probes whether the defendant's acts and statements were calculated to overcome the will of a reasonable woman. This standard . . . obviously allows men greater freedom than the second approach. It tolerates their exploitation of naive and gullible women by claiming that, in their "reasonableness calculation," the tactics should not have been threatening enough. Even at its best, the "reasonably calculated" standard creates something of a paradox: If most women have a different understanding of force than most men, then the reasonable calculation standard is one that asks how a reasonable man understands the mind of a reasonable woman. But at least it focuses primarily on the defendant's actions and thoughts and makes his guilt or innocence the center of the trial.

The final approach doesn't even do that. It judges the woman, not the man. It asks—as did the court in each of these cases—whether the will of the reasonable woman would have been overcome given the circumstances. The focus is on women generally, and on the victim as she compares (poorly) to the court's assessment of the reasonable woman. The court then proceeds to conclude that a reasonable woman's will would not have been overcome in those circumstances, because there is no "force" as men understand it.

Such an approach accomplishes two things. First, it ensures broad male freedom to "seduce" women who feel powerless, vulnerable, and afraid; the force standard guarantees men freedom to intimidate women and exploit their weaknesses, as long as they don't "fight" with them. Second, it makes clear that the responsibility and blame for such seductions belong with the woman. Because the will of a reasonable woman by definition would not have been overcome, a particular woman's submission can only mean that she is sub-par as women go or that she was complicitous in the intercourse. . . .

CONSENT

This Section will examine what has long been viewed as the most important concept in rape law—the notion of female consent. . . .

The justification for the central role of consent in the law of rape is that it protects women's choice and women's autonomy in sexual relations. Or, as one leading commentator put it: "In all cases the law of rape protects the woman's discretion by proscribing coitus contrary to her wishes." Not exactly. As discussed in the preceding Section, the law does not protect the woman from "coitus contrary to her wishes" when there is no "force." Secondly, the definition of nonconsent requires victims of rape, unlike victims of any other crime, to demonstrate their "wishes" through physical resistance. . . .

The 1981 case of *State v. Lima*[6] is . . . an example of the sort of rape case that is causing concern and division in the appeals courts today. . . .

Lima . . . involved an alleged rape in an open field involving no weapons and two individuals who knew each other. In *Lima,* the defendant had agreed to give his wife's 14-year-old cousin a ride home, and stopped at a park on the way. Once there, he pinned her shoulder to the ground with his left hand and moved his right hand inside of her blouse to her breast. When she told him to remove his hand, he told her to "shut up" and began to unbutton her shorts. She protested, "Willy, why are you doing this to me, you're my cousin," and began to cry.

Both the court of appeals and the supreme court viewed the key issue in *Lima* as whether, on these facts, the prosecution had established the "earnest resistance" of the victim required by Hawaii law. The court of appeals answered the question in the negative, because the "only resistance shown by the record are the victim's pleas to appellant to stop and an attempt to push appellant off of her." In reversing, the Supreme Court of Hawaii made much of this "push" in its interpretation of nonconsent. . . . The court affirmed the conviction on the ground that the victim "did not simply lie supine and unresisting while the respondent had his way with her." Thus, it could not be said, as a matter of law, that "the complainant here did not exhibit a 'genuine physical effort' to resist."[7] Presumably, in Hawaii, to say no, cry, and then "simply lie supine and unresisting while [a man] ha[s] his way" is to consent to sex.

Hawaii is not unique in holding that a victim must do more than say no, at least in the absence of deadly force. In *Goldberg v. State,*[8] where the defendant brought a would-be modeling prospect to his fictitious and deserted "temporary studio," his conviction of rape was reversed both on the ground that the force used was insufficient and on the ground that the victim had failed to offer "real resistance." On the latter point, the court drew a bright line between verbal and physical resistance: "It is true that she *told* the appellant she 'didn't want to do that [stuff].' But the resistance that must be shown involves not merely verbal but *physical* resistance 'to the extent of her ability at the time.' " . . .

[T]he law puts a special burden on the rape victim to prove through her actions her nonconsent (or at least to account for why her actions did not demonstrate "nonconsent"), while imposing no similar burden on the victim of trespass, battery, or robbery. . . .

Were the purpose of the consent requirement really to afford autonomy to women, there is no reason why a simple but clearly stated "no" would not suffice to signify nonconsent. Viewing women as autonomous human beings would mean treating them as persons who know what they want and mean what they say. A woman who wanted sex would say yes; a woman who did not would say no, and those verbal signals would be respected.

From a woman's point of view, the danger in this position is that many women who say "yes" are not in fact choosing freely, but are submitting because they feel a lack of power to say "no." From some men's point of view, the problem is that some women who say "no" would be willing to say "yes," or at least to "go along," if properly pressured. The "no means yes" philosophy, from this perspective, affords sexual enjoyment to those women who desire it but will not say so—at the cost of violating the integrity of all those women who say "no" and mean it.

A system of law that has traditionally celebrated female chastity and frowned upon sex outside of marriage might be expected to err on the side of less sex and to presume nonconsent in the absence

of evidence to the contrary. But if ours has now become a society in which women have been "liberated" to say yes, that provides all the more reason—if more were needed—to respect a no. If the stigma attached to saying yes has been eliminated, then so have the grounds for claiming that no means yes. . . .

[T]he consent standard could be viewed as a means to afford women their deserved freedom to engage in sex however they choose. . . . Indeed, a consent standard that allowed the individual woman to say "yes" as well as "no," to define all of the limits of permissible sex for herself and then to have that definition incorporated and respected in law, would be a means of empowering women. It could also expand liability for criminal sex to any man who refuses to respect those limits. . . .

The refusal of the law (and the society it reflects) . . . to empower women through the consent standard creates the fair warning problem that demands a resistance standard as its answer. . . . [W]e have created a fair warning problem whose only solution, at least to many courts and commentators, is to interpret force and consent in ways that punish women for "complicity" in sex, making their conduct the determinant of liability and the subject of our verdicts. There is nothing inevitable about either the problem or the solution.

NOTES

1. M. Hale, *The History of the Pleas of the Crown* 635 (1778). This statement is the usual basis for the "cautionary" instructions traditionally given in rape cases.
2. Mistakes of fact which are unrelated to elements of the offense are irrelevant to guilt or innocence; those which exculpate do so precisely because they negate the required *mens rea* as to an element of the offense. Thus, it matters not at all if the defendant believed—reasonably or unreasonably—that his victim was a professional model (the example is suggested by one case in which the defendant's stated intent was to have intercourse with a model; the victim was a clerical employee of a modeling agency); it should matter if he believes she is consenting, however, because nonconsent is an element of the crime. In order for the prohibited act to be criminal, it should not be significant, for these purposes, whether nonconsent is considered to be part of the definition of the *actus reus* prohibited by rape law, or a required "circumstance." In either case, the prosecution should be held to prove that the defendant in fact had the requisite mental state, whether it is . . . purpose, knowledge, recklessness or even negligence, as to nonconsent. See G. Williams, *Criminal Law* 137 (1953) ("It is impossible to assert that a crime requiring intention or recklessness can be committed although the accused labored under a mistake that negatived the requisite intention or recklessness. Such an assertion carries its own refutation.") . . .
3. *State v. Rush*, 289 Md. 230, 256, 424 A.2d 720, 733 (1981) (Cole, J., dissenting).
4. 310 N.C. 399, 312 S.E.2d 470 (1984).
5. *State v. Lester*, 70 N.C. App. 757, 321 S.E.2d 166 (1984), *aff'd*, 313 N.C. 595, 330 S.E.2d 205 (1985).
6. 2 Hawaii App. 18, 624 P.2d 1374 (1981), *rev'd.*, 64 Hawaii 470, 643 P.2d 536 (1982).
7. 643 P.2d at 541. . . .
8. 41 Md. App. 58, 395, A.2d 1212 (1979). . . .

REVIEW AND DISCUSSION QUESTIONS

1. Explain what Estrich sees as the basic dilemma of rape law. In what way does this dilemma reflect our society's view of sex?
2. American courts dealing with rape pay little attention to *mens rea*. Why have the results of this not been favorable to prosecutors and victims?
3. Assess the four reasons Estrich gives for her view that nonconsent as resistance is not a satisfactory substitute for *mens rea*.
4. Should a man who sincerely but falsely believes that a woman is consenting to sex be found guilty of rape? Does it make a difference whether his false belief is reasonable or unreasonable?
5. What is the "schoolboy" understanding of force, and how is it evident in the law of rape? What sort of force is necessary to constitute the crime of rape?
6. Explain the decisions in the *Alston* and *Lester* cases. What do they reveal about rape law? What, if anything, is deficient about the courts' understanding of force in these cases? According to Estrich,

what four approaches might be used to determine whether sex in threatening situations constitutes rape? What are the two implications of the approach actually adopted by the courts in these cases?

7. Nonconsent is a required element of some other crimes besides rape. What is different about the role of consent in rape? What point do the *Lima* and *Goldberg* cases illustrate?
8. Should the law of rape be reformed? If so, how?

Is the Insanity Test Insane?

R. J. Gerber

The insanity defense is a source of continuing controversy. In this essay, R.J. Gerber critically reviews the main tests of insanity that the courts have devised: the McNaughten test, the irresistible impulse test, the Durham test, and the Brawner test. All are unsatisfactory. After contending that insanity is not really a *mens rea* issue at all, Gerber argues that the *mens rea* requirement in criminal law should not be eliminated and puts forward a proposal for restructuring the legal treatment of insanity that would avoid the problems with the preceding tests of insanity.

[N]othing so preoccupies the arena of criminal law today as the fate of the insanity defense. Crime-conscious reactionaries . . . propose to inter the defense forever to curb its "unconscionable abuse." Liberal-minded reformers want to enlarge the categories of mental illness to avoid results such as *State v. Everett* where severe emotional trauma was of no avail in the sanity determination of a person who could only intellectually perceive right from wrong. Finally, some neoreactionaries of a radically liberal bent strive to abolish the defense completely as the first step toward the entire abolition of *mens rea* from criminal law. . . .

At the heart of the debate lurk two central issues: should the defense be modified or abolished? and secondly, what is the impact of altering the defense on the supposition of free will presumably at the cornerstone of criminal law itself? An approach to these major issues involves an investigation of the origin and history of the insanity defense, its various formulations, and the conflict between a "free will" vs. therapeutic concept of criminal responsibility. Of particular importance is an assessment of the advantages and disadvantages of the various existing tests, as well as the respective merits involved in the extremes of abolishing or widening the defense.

1. McNAUGHTEN AND HIS SICK MIND

The modern insanity defense derives from a bizarre English trial in 1843. Daniel McNaughten, a wood turner in Glasgow, suffered the delusion that Sir Robert Peel, the British prime minister, as well as the Jesuits and the Pope, were all conspiring against him. He couldn't get at the Pope or the Jesuits, so he determined that Peel must die. McNaughten went to London with his gun, and on January 20, 1843, in a garden next to the prime minister's house, he shot and killed Peel's secretary, Edward Drummond, believing him the prime minister.

His trial developed into a battle between medical knowledge and ancient legal authority. In anticipation of the insanity defense, the prosecutor opened with a detailed discussion of the background of criminal insanity. The origins of that defense deserve detailed review here as a prelude to understanding the eventual decision in McNaughten's trial.

2. ORIGINS OF INSANITY DEFENSE

. . . Initially, insanity was not a trial defense but a tool for pardon. The English Crown used complete insanity as a reason to grant arbitrary pardons

to convicted murderers. Madmen then were not persons but "akin to beasts" because they lacked reasoning powers. This notion bred the infamous "wild brute" or "wild beast" test of insanity. By Edward I's reign (1272–1307) the concept of insanity as "the King's grace" entered the trial process, where it functioned to mitigate punishment but not guilt.

By the time of Elizabeth I (c. 1581), a legal rationale existed for excluding certain groups from criminal responsibility. Infants and the insane were "non persons" for their supposed lack of moral capacity. They were exempted from punishment since they did not comprehend the morality of their acts. . . .

Between 1812 and 1840, three major cases framed the "wild beast" formula for insanity by requiring a total absence of criminal awareness at the time of the act. In these cases defendants, charged with murder, pleaded insanity. They all suffered from severe and actual mental disorders. On the total insanity test, one of the defendants was found guilty and executed; the other two were acquitted. . . .

From this historical record, the following conclusions emerge: first, the defense of insanity has never resulted in total freedom for the acquitted defendant. Kings and courts assumed that a defendant acquitted by reason of insanity should be restrained, not released.

Secondly, one notes the overwhelming correlation between insanity tests and free will: . . . Recurring analogies to children and "wild beasts" reflect the insane's deprivation of free will stemming from the disease of the will's counterpart, the intellect. . . . Little if any thought was given to defects in the will alone which might impair conduct.

Finally, throughout this historical excursus, one cannot escape the common law's hesitancy over partial vs. total insanity. Lying at the heart of this honest confusion as to which should exonerate are at least two important if ill-defined insights: first, sanity may be not an absolute, fixed state but a degree on a continuum with insanity—a conclusion implied by the fact that "wild beasts" and children *do* usually know rudimentary, moral differences and do intend their immoral acts. Secondly, one notes the suggestion, crystallized perhaps best for a time in the thirteenth century, that insanity should be a factor in the determination not of guilt but of pardon. These two clues are important pieces in the modern insanity jigsaw puzzle.

3. THE McNAUGHTEN TEST

At McNaughten's trial in 1843, nine medical witnesses testified that McNaughten was totally insane. Dr. Isaac Ray's book on forensic psychiatry was quoted to the court to attack Lord Hale's more restricted test of criminal insanity and the English cases resting on it. The testimony was so convincing that Chief Justice Tindal nearly directed a verdict of insanity, surprisingly commenting to the jury: ". . . I cannot help remarking . . . that the whole of the medical evidence is on one side, and that there is no part of it which leaves any doubt in the mind." Instead he committed the case to the jurors who found the defendant not guilty on the ground of insanity. McNaughten was committed to Broadmoor, a mental institution, where he later died.

After McNaughten's trial, both Houses of Parliament and Queen Victoria debated whether McNaughten's acquittal presaged a dangerously liberalized state of affairs. The fifteen common law judges were required to respond to five parliamentary questions. Their answers constitute the famous McNaughten test of insanity. . . .

> To establish a defence on the ground of insanity, it must be clearly proved, that, at the time of the committing of the act, the party accused was labouring under such a defect of reason, from disease of the mind, as not to know the nature and quality of the act he was doing, or if he did know it, that he did not know he was doing what was wrong.

. . . McNaughten's unequivocal recognition of the "right and wrong" standard as an established rule and the national attention it attracted at the time make it a case of special significance. With its sharp focus on the offender's ability to know right from wrong with respect to the particular act charged, McNaughten clarified and brought order out of existing but confusing precedent and produced a distinct, workable rule from which the more modern tests for insanity have evolved. It is the formula used in the clear majority of the

states. . . . It is a formula filled with paradox, not the least of which is that its own moral and cognitional criteria should have convicted rather than acquitted its namesake.

4. ANALYSIS OF THE McNAUGHTEN TEST

Under the McNaughten Rule, the accused is not criminally responsible if a mental disease at the time of the act prevented him from knowing (a) the nature and quality of the act or (b) that it was wrong. The basic postulate of the test is capacity to follow the right course once one is able to perceive it. Its emphasis is on the cognitive capacity to understand the moral nature of acts. The assumption is that when cognition is defective the personality as a whole is so impaired that the accused cannot "know" the wrongfulness of his actions. . . .

Above all, two phrases of the McNaughten test cause major trouble: "disease of the mind" and "know." Courts have not precisely defined "disease of the mind," but this phrase clearly is not equivalent to all the various medical classifications of mental diseases. It takes its meaning instead from the rest of the test and thus becomes limited by the word "know." Noncognitive mental disorders are therefore not "diseases of the mind." . . . In effect, to qualify as a "disease of the mind," a malady must touch an accused's reflective powers so severely that he is deprived of knowledge of the nature and quality of his act, so that he did not know what he was doing was wrong.

The word "know" may become ambiguous for persons suffering from serious mental illness not directly cognitional. A sick person's knowledge is often divorced from all affect, somewhat like the knowledge that children have of propositions they can recite but cannot understand. The knowledge crucial for criminal responsibility ought to be an appreciation or awareness of legal and moral consequences, not an abstract philosophical assent to a proposition. . . . Mental illness usually does not destroy the minimal moral awareness required by the test; it may, however, destroy the capacity to use such knowledge to control behavior.

The McNaughten test literally calls for total impairment; the accused must not know at all. Thus the traditional English hallmark of "total" insanity enshrined in the test continues to require a near impossibility. Few if any persons are "total" madmen; insanity is rather a matter of degree and context. . . .

5. THE IRRESISTIBLE IMPULSE TEST

Other tests have sought to remedy McNaughten's defects. One of these, the infrequent "Irresistible Impulse" test, is well put by New Mexico's Supreme Court, which typically joins it to the McNaughten test:

> Assuming defendant's knowledge of the nature and quality of his act and his knowledge that the act is wrong, if, by reason of disease of the mind, defendant has been deprived of or has lost the power of his will which would enable him to prevent himself from doing the act, he can not be found guilty.

The test represents the first wholesale amendment of McNaughten. Text writers have misnamed this the "irresistible impulse" test; it is really a control test. Goldstein has found so many differing formulations of the test that "there is no monolith called the 'irresistible impulse' test." Their central idea, however, is loss of control, and their universal motivation is to remedy McNaughten's omission of noncognitive control data.

Weighty objections, however, cut against a narrowly formulated control test that connotes the idea of sudden, impulsive or irresistible loss of control. Just as an intellectual interpretation of the McNaughten knowledge requirement ignores emotional factors, a "sudden impulse" interpretation of the control test may restrict evidence and jury deliberations solely to immediacy, thus keeping from the jury the durational nature of an accused's mental condition. The "impulse" connotation fails to account for cases where loss of volition takes place not suddenly but gradually, as in cases of melancholia and paranoia. . . . It gives no consideration to the defendant whose mental disease is characterized by brooding, introspection and reflection. The connotation ascribed to"control" means that an individual with sufficient powers temporarily to control or forestall his impulses is legally sane even though he ultimately surrendered completely to these impulses. . . .

6. THE DURHAM TEST

In 1954, in Washington, D.C., the United States Court of Appeals for the District of Columbia Circuit reversed the conviction of a Monte Durham for housebreaking and petit larceny. The court held that Durham had presented enough evidence on the question of his insanity to raise an issue for decision under the District of Columbia's version of the McNaughten test for criminal responsibility. In the course of an extensive opinion, however, the court rejected the McNaughten rule in all its formulations and held that "a broader test" should be adopted for the District of Columbia: "[A]n accused is not criminally responsible if his unlawful act was the product of a mental disease or mental defect." . . .

Widely hailed by both jurists and psychiatrists, *Durham* seemingly represented a scientific advance in criminal jurisprudence. Clearly, it represents the psychiatrization of the criminal law, an attempt to transform into legal terms the notion that there are two modes of existence—one sane, and the other insane. The *Durham* standard views mental functioning as essentially unitary but multifaced. No single mental faculty determines the existence or nonexistence of sanity, just as no single faculty is responsible for the control of human behavior. Impaired control may result from a wide variety of causes in the psyche, not all of which are cognitional.

Yet ambiguities lingered with *Durham.* While the approach permitted the psychiatrist to testify on the full range of rational, emotional, and volitional elements influencing an individual's psyche, this test drew a single line on one side of which a defendant was held criminally responsible for his actions and on the other side of which he was completely exculpated. The test does not account for the psychic reality of a continuous graduated scale of responsibility. . . .

Defining "mental disease" became another acute problem. Early experience under *Durham* was that "mental disease" was synonymous with psychosis, with the apparent agreement of all the participants in the trial process. Psychiatrists believed that they were using a legal term with an understood legal meaning; lawyers believed they were using a medical term with a standard medical content. The problem burst into the open in 1957, when the staff of Saint Elizabeth's Hospital decided that nonpsychotic diagnosis—particularly, the diagnosis of "sociopathic personality disturbances"—would be explicitly recorded in reports of mental examinations, thus opening the insanity defense to a large number of defendants. . . .

Ultimately at least [several] intertwined problems arose with the vague "mental disease" concept in *Durham.* In the first place, *Durham* failed to define the parameters of mental disease clearly enough to determine whether sociopathy would count. Durham himself was diagnosed a sociopath and held sane. Shortly thereafter one Blocker, on trial for murdering his wife, pled insanity, was diagnosed a sociopath, and also found guilty. One month later, however, when the St. Elizabeth's staff broadened its forensic concept of mental disease to include sociopathy, Blocker appealed and was granted a new trial. This sudden reversal of psychiatric opinion on the legal process provoked judicial tempers. . . .

The first problem brought a second: the domination of the courtroom by psychiatrists. Whereas Judge Bazelon [who wrote the *Durham* opinion] merely wanted to open the door to liberal psychiatric testimony, he in fact opened the floodgates to medical conjecture frequently inscrutable and just as frequently overbearing for judges and jurors. Particularly onerous to the jury's role was the tendency by psychiatrists to make unfounded statements regarding the vexing issue of just which psychic conditions did or did not cause the criminal act, thus permitting psychiatrists to determine criminal responsibility. . . .

7. BRAWNER: A NEW TEST

. . . In the *Brawner* case, the defendant's jaw had been broken in a fight. He returned half an hour later with a gun, fired five shots through a closed apartment door, and killed one of the occupants. Four expert witnesses agreed that Brawner was suffering from a psychiatric or neurological abnormality, variously described as "epileptic personality disorder," "personality disorder associated with epilepsy." The prosecution witnesses apparently regarded Brawner's actions as a normal response in the circumstances, a more or less legitimate response. He had been severely injured in a fight and was very vindictive. As one of the doctors

testified, "I think I would, too, under the same circumstances, want to get even with somebody who broke my jaw."

Determined to abandon *Durham,* the *Brawner* court turned to the ALI's proposed rule of criminal responsibility:

> **1.** A person is not responsible for criminal conduct if at the time of such conduct as a result of mental disease or defect he lacks substantial capacity either to appreciate the criminality [wrongfulness] of his conduct or to conform his conduct to the requirements of law.
>
> **2.** As used in this Article, the terms "mental disease or defect" do not include an abnormality manifested only by repeated criminal or otherwise antisocial conduct.

. . . Critical reading of *Brawner in toto* reveals serious inconsistencies touching the insanity defense relative to *mens rea.* The court's jury instruction on insanity provides: "You are not to consider this defense unless you have first found that the government has proved beyond a reasonable doubt each essential element of the offense." This instruction obviously means that insanity becomes an issue only after a prior finding that the defendant (1) did the criminal act (2) with its defining mental element, viz., intentionally, knowingly, etc. The defense comes alive then only for defendants already inculpated in crime. As a separate issue from *mens rea,* the defense is not an evidentiary standard for weighing intent or free will at all but an escape valve for persons whose mental state, while criminal, is disturbed beyond the range of full responsibility.

Having established this commendable distinction between *mens rea* and the insanity defense, the court next erases that distinction completely in saying it refuses to abolish the insanity defense because its abolition would eliminate free will from the criminal law. The court could reach this conclusion only by confusing the proposal to abolish the insanity defense with the quite broader proposal to abolish *mens rea.* It further mistakenly assumes that the abolition of insanity would inaugurate strict liability crimes, i.e., criminality apart from criminal intent, when in fact that result follows only from the abolition of *mens rea.* These assumptions are inconsistent with the court's own suggested jury instruction. They contribute nothing toward clarifying the role of insanity relative to *mens rea.* . . .

In hopes of clearing the air for a possible solution, *Brawner* generates the need to analyze the following propositions anew: (1) insanity is not a *mens rea* issue; (2) *mens rea* should not be abolished, and (3) the insanity defense needs to be restructured.

8. INSANITY IS NOT A *MENS REA* ISSUE

Ennui over the proliferation of tests has spawned the proposal to abolish the insanity defense completely. Some critics have gone considerably further to urge the abolition of the defense not merely for its own sake but as a first step toward abolishing *mens rea* from criminal law. Before the ultimate merits of either proposal can be considered, it is imperative to realize that insanity is not properly a *mens rea* question, but a unique issue distinct from criminal intent.

Mens rea means criminal intent. It refers to the mental element which, together with a physical act, constitutes a crime. These mental elements, as defined by the Model Penal Code and most revised state codes, basically can appear in any one of three forms: (1) "Intentionally," (2) "Knowingly," or (3) "Recklessly, i.e., with disregard of known risks." Intentionally means acting toward a conscious criminal goal. Knowingly refers to acting with the awareness of the circumstances of the act. Recklessly means acting with the awareness that certain dangerous results are likely to occur. Any one of the three mental states constitutes *mens rea.*

Absence of criminal intent means that no crime has been committed, in effect, that a bodily movement occurred independently of the mind's acquiescence. The absence of *mens rea* thus establishes an innocence claim. This claim is not a defense but a denial that any crime occurred at all.

Insanity is a wholly separate matter. It is not an innocence claim denying a crime, but an excuse to an established crime. Insanity asserts that although a crime has occurred, some sort of irrationality colored the guilty criminal intent. It acts as a second wind, as a second string in the bow of an already guilty defendant. As *Brawner* confusedly perceived, insanity is not really a negation of intent or knowledge or disregard of risk but a psychic distortion coloring these guilty and quite compatible mental states. Like

the similar excuses of infancy and necessity, insanity tacitly admits both the act and the guilty criminal intent but asserts the need for some exception from the punitive consequences of guilt.

The foregoing analysis supports two conclusions: the first is that the fates of *mens rea* and insanity are separate. Altering the defense does not entail the abolition of *mens rea* or of the free will postulate upholding *mens rea.* Hence the disutility of the widespread bellowing that the insanity defense is necessary in order to retain the entire concept of free will at the heart of the criminal law system.

The second conclusion is more psychiatric and factual: as any defense lawyer knows, the vast majority if not all insane criminals do actually intend or know or recklessly risk their criminal acts. To take one example which stands for all, McNaughten himself quite obviously intended the death of the prime minister. He knew what he was doing, wanted its result, and knew that it was frowned upon by his peers. The same can be said for every or nearly every insane defendant. It is simply not the case that the insane are incapable of intention, deliberation, moral understanding etc.; rather, like "wild beasts" and children, their mental operations occur within an unusually restricted context of a certain few misinterpreted objects, the context of which prevents the usual inhibitions from arresting the criminal plan. One may say, of course, that the insane defendant didn't "fully" or "really" know what he was doing; but such an interpretation cannot cloud the fact that the insane's knowledge and intention nonetheless are notably at work in the planning of the criminal act, often to the point of being overbearing. In a word, insanity rarely if ever precludes a criminal *mens rea.*

If the insane do in fact intend and know their criminal acts, such evidence properly touches on a finding of *mens rea,* not of insanity. Hence, their fate on the guilt-innocence issue should be dependent upon the much debated fate of *mens rea,* not on the role of insanity.

9. *MENS REA* CANNOT BE ABOLISHED

The criminological positivists in the early 1900s first proposed to eliminate from the criminal law the entire *mens rea* requirement. While complex, their motivation was one of social defense which placed the security of society above the fate of the individual. In more recent years this proposal has taken a therapeutic pallor. Lady Wootten argues that *mens rea* must go because we cannot get inside a person's mind to really see intent at work.[1] Karl Menninger feels that all crime is really sickness, so there is no point in showing a sick mind in a criminal trial.[2] Kittrie urges the abolition of *mens rea* as the essential step toward a "*parens patriae* therapeutic state."[3] Their proposals are not directed precisely at insanity; they seek to eliminate *all* evidence of mental state from the criminal law so as to replace the concepts of guilt, blame, free will, and responsibility with the simpler concept of social danger.

Packer suggests several of the likely assumptions that underlie this therapeutic system of criminal law:

> **1.** Free will is an illusion because human conduct is determined by forces that lie beyond the power of the individual to modify.
>
> **2.** Moral responsibility, accordingly, is an illusion because blame cannot be ascribed for behavior that is ineluctably conditioned.
>
> **3.** Human conduct, being causally determined, can and should be scientifically studied and controlled.
>
> **4.** The function of the criminal law should be purely and simply to bring into play processes for modifying the personality, and hence the behavior of people who commit them in the future; or, if all else fails, to restrain them from committing offenses by the use of external compulsion (e.g., confinement).
>
> **5.** We have . . . real knowledge about how to rehabilitate people.
>
> **6.** We know how to predict those who exhibit traits that are dangerous.[4]

The advocates of this streamlined criminal law see it as more scientific, manageable, and forward-looking than the existing punishment-oriented free will system rooted in the moral concept of *mens rea.*

There are, however, any number of cogent reasons why *mens rea* cannot be abolished. At the phenomenological level, *mens rea* is simply a legalistic way of enshrining (perhaps confusing) certain patent nonlegal concepts that operate effectively in everyday noncriminal contexts. *Mens rea* is a

crystallization of a pervasive but elusive rationale of responsibility based on assessment of purpose. In the everyday world, people intend goals, direct means to ends, and anticipate consequences of their acts. When a person is shoved in a crowd, for example, the "shovee's" response to the "shover" depends entirely upon whether the shove is interpreted as deliberate or accidental.

Similarly, punishment and administrative consequences differ for a person who causes an automobile accident through recklessness as opposed to a person whose automobile causes the same kind of accident through some noncontrollable mechanical defect. Within the criminal law, *mens rea* adopts these everyday assumptions of intent for evaluating persons charged with crime. *Mens rea* thus is rooted not so much in legal history as in human history, in the everyday working predication of merit and blame on those who are deserving because they intend their acts. A public trial is a public morality play that serves to underscore these everyday concepts of intentionality and knowledge as valid criteria for evaluating behavior. *Mens rea* is actually an assertion that human society is a society of persons who do not view themselves as mere bodies moving as harmful vectors; persons interpret the movements of other people as manifestations of intention and choice. We can and do read the intentionality of persons in and through their words and deeds. Abolition of *mens rea* from the law would not abolish it from the everyday world.

To this phenomenological analysis must be added a further, more psychological objection: abolition of *mens rea* makes for strict liability for all crimes, with the crimes themselves reduced to merely bodily acts. Manslaughter becomes undistinguishable from homicide, tax fraud no different than an error in arithmetic (and, of course, vice versa). Inchoate crimes like attempt and conspiracy become eliminated or, more likely, overextended. Many crimes would merge with torts; assault, battery, false imprisonment would be undistinguishable as crimes or torts, with the curious result that criminals might be subject to dollar damages and ordinary tortfeasors or accident-prone persons subject to prison or other criminal terms. Wootten's strict liability coalesces exactly with Skinner's notion that it is not the person but the environment which needs the blame—and the change. Crime is in the woodwork, not in the criminal. Every error becomes a crime.

To this Clockwork Orange scenario are two further, obvious objections: the emotional one is that such a degrading picture of human responsibility will never appeal to more than mad academics. The second is that the public experiences the "free will myth" in certain central acts like long-range planning and resistance to negative conditioning. This experience, valid or not, supports the assurance of free will resting at the heart of *mens rea.* Even if this is illusion, it is an illusion that all persons experience and which the criminal law necessarily reinforces.

The Wootten-Skinner-Menninger-Kittrie proposal to abolish *mens rea* is impractical for a final reason. It substitutes for "bad" and "mad" the nebulous concept of social deviancy, which would subject to state control not only the bad and the mad but all tort-feasors, idealist reformers and eccentrics—in a word, all nonconformists. This kind of law encroaches on all, for once guilt is abolished, so too is innocence. The trouble with a criminal policy based on social defense is that the individual has no defense against an infallible, static society. Such people as Martin Luther King, Daniel Berrigan, William Sloan Coffin and Henry David Thoreau need "treatments," if at all, differing in degree and in kind from that imposed on the violent armed robber and rapist. With the abolition of *mens rea,* criminal law loses any discriminating index of culpability and hence of disposition. Despite Wootten's deterministic assumption that crime is a unilateral bodily concept, the fact is that a mental chasm separates a genetic compulsion to crime (such as the XYY chromosome) from deliberate, prolonged defiance to a detested law (as in the instance of war protestors or civil rights activists). And neither of these *mens rea* covers a third, more commonplace situation where crime reflects a spontaneous attempt to capitalize on an immediate opportunity (as in most burglaries and auto thefts). The point is that mental state varies for varying crimes and its variation dictates varying degrees of culpability. If mental state becomes either a unilateral or a nonexistent concept in defining criminality, neither the severity of the crime nor the propriety of treatment can be determined.

10. INSANITY NEEDS RESTRUCTURING

Retaining *mens rea* need not blind one, as it did the *Brawner* court, to the problems involved in retaining any or all of the insanity tests. Regardless of form, each test has at least five lingering problems: (1) exclusion of certain evidence; (2) domination of the lay jurors by medical experts; (3) discrediting psychiatric testimony in the adversary process; (4) the conception of sanity in absolute rather than relative terms of degree on a continuum; and (5) the eventual inconsistency of detaining a person found not guilty by reason of insanity.

To solve these . . . problems . . . [we should] consider seriously the propriety of a two-phase court procedure in cases where sanity is at issue. Phase one would be a conventional criminal trial to determine two precise things: (1) whether the defendant did the act, and (2) whether he did it with its defining mental state. Affirmative findings on both of these by judge or jury would result in a guilty verdict. The only other alternative would be outright acquittal. . . . [I]nsanity evidence would be admissible only to the extent that it negates the precise *mens rea* defining the crime, which in modernized codes strictly means only to the extent it negates intentionality, knowledge, or consciousness of risk (or whatever other precise mental state is defined in the description of the offense). In practice, this means that very few insane defendants could successfully introduce evidence of insanity in this trial. Certainly, McNaughten himself could not, nor could Jack Ruby nor Lee Harvey Oswald, nor most truly insane defendants, since insanity is quite compatible with intent, moral awareness, and awareness of risk. Furthermore, where *mens rea* is really at issue, the kind of evidence adduced is not likely to be medical diagnoses but circumstantial, factual data touching on lack of planning, scheme, design, lack of awareness, etc. Except for the rare case where insanity may negate the required mental state, psychiatric jargon would not be introduced into evidence before a jury.

Following a guilty verdict, a second procedure would become available at the option of the defense: a court-ordered psychiatric test of the defendant by appropriate psychiatrists followed by a sanity hearing—a nonadversary procedure before a judge, not a jury. The narrow purpose of this hearing would be for evaluating evidence of reduced responsibility, including perhaps not only insanity but diminished responsibility or voluntariness generally. Unlike the trial, this hearing would determine if defendant's sanity is or was reduced to such a degree that it would be unjust to imprison him when other remedies are more likely to protect society and to assure his beneficial return to it. In practice, the dispositional alternatives would be not only prison but also hospitals and medical or community-related therapy programs. Most importantly, at the termination of this hearing, the judge would make a finding on the defendant's degree of sanity. . . . If the court found the defendant insane, the judge thereupon would amend the original jury verdict of guilt to "guilty of the act but insane." This amended verdict would be his mandate for committing the defendant to a treatment program rather than to prison. . . .

The advantages of such a proposal are obvious. In the first place, insanity and guilt would be retained as major factors in criminal law. Insanity would operate, however, not in the context of disputes over the wording of ambiguous medical tests, not in the context of the adversary guilt-innocence issue, but rather in the early common law tradition as a mitigating factor decisive in pardon and disposition. Precise insanity tests would be unnecessary at both the trial and the sanity hearing. . . . Criminal trials would speed up. Jurors would decide guilt but not medically mitigating factors. Prosecutors would not be encumbered with insanity defenses threatening loss of their cases. Defendants would possess, at their option, a second exculpatory string in their bows to mitigate guilt and to receive medical care. The truly insane would receive neither guilt nor punishment in the conventional sense. Such a procedure would remedy the absolutism of present insanity tests, which now envision insanity-sanity in all-or-nothing terms, when in fact sanity is a matter of fluctuating degree. Finally, this procedure would extend the use of insanity evidence to a far broader class than the predominantly serious felons who now invoke it, with the expectation that insane car thieves and burglars might get, at their option, psychiatric rather than purely penal attention, without having to confuse a reluctant jury to do so.

NOTES

1. Wootten, *Crime and the Criminal Law* (1963).
2. Menninger, "Medico-Legal Proposals of the American Psychiatric Association," *J. Crim. L. C. and P. S.*, 19 (1928), pp. 367, 373. Lately Menninger has tempered his view to the confusing point where he now says crime "is not an illness, although I think it should be," in *The Crime of Punishment* (1968), p. 254.
3. Kittrie's views are expressed best in "The Therapeutic State," *Wayne L. R.* 19 (1975), p. 871.
4. Packer, *The Limits of The Criminal Sanction* (1969). Cf. also his "Enemies of Progress," *New York Review of Books*, Oct. 23, 1969, p. 17.

REVIEW AND DISCUSSION QUESTIONS

1. What is the McNaughten test, and what are the problems with it?
2. Assess the strong and weak points of the irresistible impulse, Durham, and Brawner tests. Which is the soundest?
3. On what grounds does Gerber contend that insanity is not a *mens rea* issue? Do you agree? Compare Gerber's understanding of *mens rea* with that of Brandt.
4. What reasons can be given for abolishing *mens rea?* Do you agree with Gerber that it should not be abolished?
5. Assess Gerber's proposal for restructuring the legal treatment of insanity. Do you see any difficulties with it? Is the issue of insanity sometimes relevant to guilt? Does Gerber's proposal avoid the problems he has raised for the courts' various tests of insanity?
6. "Gerber's proposal would allow criminals to avoid punishment." Assess this statement, explaining why you agree or disagree with it.

What Is So Special About Mental Illness?

Joel Feinberg

Many legal writers argue that mental illness is not itself an independent basis of legal exculpation; at most, it is evidence that one of the traditional excuses—compulsion or ignorance—applies. After explaining this view and why it is plausible, Joel Feinberg, who was professor of philosophy at the University of Arizona, raises some doubts about it. In particular, the standard view does not do full justice to the moral significance of mental illness as it bears on blame and punishment. At the heart of criminal behavior that is both voluntary and sick, Feinberg argues, is the irrationality of the criminal's motives. While punishment in such case may be justified to deter others, it does not seem to be an appropriate way of dealing with the offender himself.

Professor Dershowitz has [argued that] . . . mental illness should not itself be an independent ground of exculpation, but only a sign that one of the traditional standard grounds—compulsion, ignorance of fact, or excusable ignorance of law—may apply.[1] Mental illness, then, while often relevant to questions of responsibility, is no more significant—and significant in no different way—than other sources of compulsion and misapprehension.

Now although I am almost completely convinced that this is the correct account of the matter, I am nevertheless going to air my few lingering doubts as if they were potent objections, just to see what will happen to them. I shall suggest, then, in what follows,

that mental illness has an independent significance for questions of responsibility not fully accounted for by reference to its power to deprive one of the capacity to be law-abiding.

1

At the outset we must distinguish two questions about the relation of mental illness to criminal punishment. . . .

> **1.** How are mentally sick persons to be distinguished from normal persons?
>
> **2.** When should we accept mental illness as an excuse?

The first appears to be a medical question that requires the expertise of the psychiatrist to answer; the second appears to be an essentially controversial question of public policy that cannot be answered by referring to the special expertise of any particular group.

Some psychiatrists may wish to deny this rigid separation between the two questions. They might hold it self-evident that sick people are not to be treated as responsible people; hence the criteria of illness are themselves criteria of nonresponsibility. But, obviously, this won't do. First of all, the fact of illness itself, even greatly incapacitating illness, does not automatically lead us to withhold ascription of responsibility, or else we would treat *physical* illness as an automatic excuse. But in fact we would not change our judgments of Bonnie and Clyde one jot if we discovered that they both had had 103-degree fevers during one of their bank robberies, or of Al Capone if we learned that he had ordered one of his gangland assassinations while suffering from an advanced case of chicken pox. Secondly, there are various crimes that can be committed by persons suffering from mental illness that can have no relevant bearing on their motivation. We may take exhibitionism to be an excuse for indecent exposure, or pedophilia for child molestation, but neither would be a plausible defense to the charge of income-tax evasion or price-fixing conspiracy. These examples show, I think, that the mere fact of mental illness, no more than the mere fact of physical illness, automatically excuses. We need some further criterion, then, for distinguishing cases of mental illness that do excuse from those that do not, and this further question is not an exclusively psychiatric one. What we want to know is this: what is it about mental illness that makes it an excuse when it is an excuse?

So much, I think, is clear. But now there are two types of moves open to us. The first is preferred by most legal writers, and it is the one about which I intend to raise some doubts. According to this view, there is nothing very special about mental disease as such. Mental illness is only one of numerous possible causes of *incapacity,* and it is incapacity—or, more precisely, the incapability to conform to law—that is incompatible with responsibility. Ultimately, there is only one kind of consideration that should lead us to exempt a person from responsibility for his wrongful deeds, and that is that he *couldn't help it.* Sometimes a mental illness compels a man to do wrong, or at least makes it unreasonably difficult for him to abstain. . . . But in other cases, as we have seen, mental illness no more compels a given wrongful act than the chicken pox does, or may be totally irrelevant to the explanation of the wrongdoing. . . . What counts, then, for questions of responsibility is whether the accused could have helped himself, not whether he was mentally well or ill.

Aristotle put much the same point in somewhat different but equally familiar language. A man is responsible, said Aristotle, for all and only those of his actions that were voluntary; to whatever extent we think a given action less than voluntary, to that extent we are inclined to exempt the actor from responsibility for it. There are, according to Aristotle, two primary ways in which an action can fail to be voluntary: it can be the result of *compulsion,* or it can be done in *ignorance.* Thus if a hurricane wind blows you twenty yards across a street, you cannot be said to have crossed the street voluntarily, since you were compelled to do it and given no choice at all in the matter. And if you put arsenic in your wife's coffee honestly but mistakenly believing it to be sugar, you cannot be said to have poisoned her voluntarily, since you acted in genuine ignorance of what you were doing.

. . . On the view I am considering . . . the mental illness of an actor is not still a third way in which his

actions might fail to be voluntary; rather, it is a factor which may or may not compel him to act in certain ways, or which may or may not delude, or mislead, or misinform him in ways that would lead him to act in ignorance. Indeed, on this view, mental illness ought not even to be an independent category of exculpation on a level with, say, self-defense or mistake of fact. Self-defense and relevant blameless mistakes of fact always excuse, whereas mental illness excuses only when it compels or deludes. . . .

The nineteenth-century judges who formulated the famous McNaghten Rules were presumably quite sympathetic with the view I have been describing, that there really is nothing very special about mental illness. These rules are not at all concerned with neurotically compulsive behavior—a category which simply was not before their minds at the time. Rather, they were concerned with those dramatic and conspicuous disorders that involve what we call today "paranoid delusions" and "psychotic hallucinations." The interesting thing about the rules is that they treat these aberrations precisely the same as any other innocent "mistakes of fact"; in effect the main point of this part of the McNaghten Rules is to acknowledge that mistakes of fact resulting from "disease of the mind" really are genuine and innocent and, therefore, have the same exculpatory force as more commonplace errors and false beliefs. The rules state that, "when a man acts under an insane delusion, then he is excused only when it is the case that *if* the facts were as he supposed them his act would be innocent." . . . Thus if a man suffers the insane delusion that a passerby on the street is an enemy agent about to launch a mortal attack on him and kills him in what he thinks is "self-defense," he is excused, since if the facts were as he falsely supposed them to be, his act would have been innocent. But if . . . he shoots his wife because, in his insane delusion, he thinks her hair has turned gray, that would not have been an allowable defense. Note that the mental disease that leads to the insane delusion in these instances is given no special significance except insofar as it mediates the application of another kind of defense that can be used by mentally healthy as well as mentally ill defendants. . . .

To summarize the view I have been considering: a mere finding of mental illness is not itself a sufficient ground for exempting a person from responsibility for a given action; nor is it a sufficient ground for finding him not to be a responsible or competent person generally, with the loss of civil rights such a finding necessarily entails. At most, in criminal proceedings mental illness may be evidence that one of the traditional grounds for moral exculpation—compulsion or ignorance—applies to the case at hand, and in civil commitment hearings it may be evidence of dangerousness or of cognitive impairment. But it has no independent moral or legal significance in itself either as an excuse or as a ground for commitment.

2

. . . But . . . I have a lingering doubt that the above account does full justice to the moral significance of mental illness as it bears on blame and punishment. I shall devote the remainder of my remarks to a statement of that doubt.

Let me turn immediately to the kind of case that troubles me. I have in mind cases of criminal conduct which appear to be both voluntary (by the usual Aristotelian tests) and sick. Let me give some examples and then contrast them with normal voluntary criminal acts.

First consider a nonviolent child molester. He is sexually attracted to five- and six-year-old boys and girls. His rational faculties are perfectly normal. He knows that sexual contacts with children are forbidden by the criminal law, and he takes no unnecessary risks of detection. For the most part, he manages to do without sex altogether. When he does molest a child he characteristically feels guilt, if not remorse, afterward. He has no understanding of his own motivation and often regrets that his tastes are so odd.

Next in our rogues' gallery is a repetitive exhibitionist. He has been arrested numerous times for exposing his genitalia in public. He does this not to solicit or threaten, but simply to derive satisfaction from the act itself: exposure for exposure's sake. For some reason he cannot understand, he finds such exposure immensely gratifying. Still, he knows that it is offensive to others, that it is in a way publicly humiliating, that it is prohibited by law, and that the chances of being caught and punished are always very great. These things trouble him much, and they

often, but not always, lead him to restrain himself when the impulse to self-exposure arises.

My third example is drawn from a landmark case in the criminal law, one of the first in which kleptomania was accepted as an excuse: *State v. McCullough,* 114 Iowa 532 (1901). The defendant, a high school student, was charged with stealing a school book worth seventy-five cents. It was discovered that stolen property in his possession included "14 silverine watches, 2 old brass watches, 2 old clocks, 24 razors, 21 pairs of cuff buttons, 15 watch chains, 6 pistols, 7 combs, 34 jack knives, 9 bicycle wrenches, 4 padlocks, 7 pair of clippers, 3 bicycle saddles, 1 box of old keys, 4 pairs of scissors, 5 pocket mirrors, 6 mouth organs, rulers, bolts, calipers, oil cans, washers, punches, pulleys, spoons, penholders, ramrods, violin strings, etc." ["etc."!] One can barely imagine the great price in anxiety this boy must have paid for his vast accumulation of worthless junk.

Finally, consider a well-off man who shoplifts only one kind of item, women's brassieres. He could easily afford to pay for these items and, indeed, often does when there is no other way of getting them, or when he is in danger of being caught. He does not enjoy stealing them and suffers great anxiety in worrying about being found out. Yet his storerooms are overflowing with brassieres. He burgles homes only to steal them; he assaults women only to rip off their brassieres and flee. And if you ask him for an explanation of his bizarre conduct, he will confess himself as puzzled by it as any observer.

Now, for contrast, consider some typical voluntary normal crimes. A respectable middle-aged bank teller, after weighing the risks carefully, embezzles bank funds and runs off to Mexico with his expensive lady friend. A homeowner in desperate need of cash sets his own house on fire to defraud an insurance company. A teenager steals a parked car and drives to a nearby city for a thrill. An angry man consumed with jealousy, or indignation, or vengefulness, or spite, commits criminal battery on a person he hates. A revolutionary throws a bomb at the king's carriage during an insurrection. These criminals act from a great variety of unmysterious motives—avarice, gain, lust, hate, ideological zeal; they are all rationally capable of calculating risks; they all act voluntarily.

How do the "sick" criminals in my earlier list differ from these normal ones? We might be tempted to answer that the pedophiliac, the exhibitionist, the kleptomaniac, and the fetishist are all "compulsives" and that their criminal conduct is therefore not entirely voluntary after all; but I believe it is important to understand that this answer is unsatisfactory. There is no *a priori* reason why the desires, impulses, and motives that lead a person to do bizarre things need necessarily be more powerful or compulsive than the desires that lead normal men to do perfectly ordinary things. It is by no means self-evident, for example, that the sex drives of a pedophiliac, an exhibitionist, or a homosexual must always be stronger than the sexual desires normal men and women feel for one another.

There is much obscurity in the notion of the "strength of a desire," but I think several points are clear and relevant to our purposes. The first is that, strictly speaking, no impulse is "irresistible." For every case of giving in to a desire, I would argue, it will be true that, if the person had tried harder, he would have resisted it successfully. The psychological situation is never—or hardly ever—like that of the man who hangs from a windowsill by his fingernails until the sheer physical force of gravity rips his nails off and sends him plummeting to the ground, or like that of the man who dives from a sinking ship in the middle of the ocean and swims until he is exhausted and then drowns. Human endurance puts a severe limit on how long one can stay afloat in an ocean; but there is no comparable limit to our ability to resist temptation. Nevertheless, it does make sense to say that some desires are stronger than others and that some have an intensity and power that are felt as overwhelming. Some desires, in fact, may be so difficult to resist for a given person in a given state at a given time that it would be unreasonable to expect him to resist. A dieting man with a strong sweet tooth may find it difficult to resist eating an ice cream sundae for dessert; but a man who has not eaten for a week will have a much harder time still resisting the desire to eat a loaf of bread, which just happens to belong to his neighbor. Any person in a weakened condition, whether the cause be hunger or depression, fatigue or gripping emotion, will be less able to resist any given antisocial impulse than

a person in a normal condition. But, again, there is no reason to suppose that bizarre appetites and odd tastes are always connected with a "weakened condition," so that they are necessarily more difficult to resist than ordinary desires. And thus there is no reason to suppose that so-called sick desires must always be compulsive or unreasonably difficult to resist.

It might seem to follow that there is *no* morally significant difference between normal and mentally ill offenders, that the one class is just as responsible as the other, provided only that their criminal actions are voluntary in the usual sense. But if this is the proper conclusion, then I am at a loss to see what difference there can be between mental illness and plain wickedness. As an ordinary citizen, before I begin to get confused by philosophy, I sometimes permit myself to feel anger and outrage at normal criminals, whereas I cannot help feeling some pity (mixed, perhaps, with repugnance) toward those whose conduct appears bizarre and unnatural. But unless I can find some morally telling difference between the two classes of criminals, then these natural attitudes must be radically reshaped, so that the fetish thief, for example, be thought as wicked as the professional burglar.

There do seem to be some striking differences between the two classes, however, and perhaps some of these can rescue my prephilosophical attitudes. Most of them have to do not with the criminal's intentions, but with his underlying motivation—the basis of the appeal in his immediate goals or objectives. The first such difference is that the sick criminal's motives appear quite *unintelligible* to us. We sometimes express our puzzlement by saying that his crimes have no apparent motive at all. We cannot see any better than the criminal himself "what he gets out of it," and it overburdens our imaginative faculties to put ourselves in his shoes. We understand the avaricious, irascible, or jealous man's motives all too well, and we resent him for them. But where crimes resist explanation in terms of ordinary motives, we hardly know what to resent. Here the old maxim "to understand all is to forgive all" seems to be turned on its ear. It is closer to the truth to say of mentally ill wrongdoers that to forgive is to despair of understanding.

Yet mere unintelligibility of motive is not likely to advance our search for the moral significance of mental illness very far, especially if we take the criterion of unintelligibility in turn to be the frustration of our "imaginative capacities" to put ourselves in the criminal's shoes and understand what he gets out of his crimes. This test of imaginability is far too elastic and variable. On the one hand, it seems too loose, since it permits the classification as unintelligible (or even sick) of *any* particular passion or taste, provided only that it is sufficiently different from those of the person making the judgment. Some nonsmokers cannot understand what smokers get out of their noxious habit, and males can hardly understand what it is like to enjoy bearing children. On the other hand, once we begin tightening up the test of imaginability, there is likely to be no stopping place short of the point at which *all* motives become intelligible to anyone with a moderately good imagination and sense of analogy. The important thing is not that the sick criminal's motives may seem unintelligible, but rather that they are unintelligible in a certain respect or for a certain reason.

We get closer to the heart of the matter, I think, if we say that the mentally ill criminal's motives are unintelligible because they are irrational—not just unreasonable, but *irrational.* All voluntary wrongdoing, of course, is unreasonable. It is always unreasonable conduct to promote one's own good at another's expense, to be cruel, deceitful, or unfair. But in a proper sense of "rational," made familiar by economists and lawyers, wrongdoing, though unreasonable, can be perfectly rational. A wrongdoer might well calculate his own interests, and gains and risks thereto, and decide to advance them at another's expense, without making a single intellectual mistake. A rational motive, in the present sense, is simply a *self-interested* motive, or perhaps an intelligently self-interested one. The motives of mentally ill criminals are not usually very self-interested. The Supreme Court of Iowa, in overturning the conviction of young McCullough, held that the question of the jury should have been: did the accused steal because of a mental disease driving him by "an insane and irresistible impulse" to steal, or did he commit said acts "through excessive greed or avarice?" The Court's alternatives are not exhaustive. Very likely McCullough's impulses were neither irresistible nor "greedy and avaricious." Greed and avarice

are forms of selfishness, excessive desires for material goods and riches for oneself. As motives they are preeminently self-interested and "rational." McCullough's sick desires, however, were not for his own good, material or otherwise. He stole objects that could do him no good at all and assumed irrational risks in the process. The desire to steal and hoard these useless trinkets was a genuine enough desire, and it was *his* desire; but it does not follow that it was a desire to promote his own good.

This point too, however, can be overstated. It may well be true that none of the mentally ill crimes we are considering is done from a self-interested motive, but this feature hardly distinguishes them (yet) from a wide variety of voluntary crimes of great blameworthiness committed by perfectly normal criminals. By no means all voluntary crimes by normal criminals are done from the motive of gain. Some are done to advance or retard a cause, to help a loved one, or to hurt an enemy, often at great cost to the criminal's self-interest. What distinguishes the sick crimes we have been considering is not that they are unself-interested, but rather that they are *not interested at all.* They do not further *any* of the actor's interests, self *or* other-regarding, benevolent or malevolent. The fetishist's shop-lifting is not rational and self-serving; he attains no economic objective by it. But neither does it hurt anyone he hates nor help anyone he loves; it neither gains him good will and prestige, nor satisfies his conscience, nor fulfills his ideals. It is, in short, not interested behavior.

But even this distinction does not quite get to the very core of the matter. The fetishist's behavior not only fails to be interested; it fails even to appear interested to him. To be sure, it is designed to fulfill the desire which is its immediate motive; but fulfillment of desire is not necessarily the same thing as abiding satisfaction. He may be gratified or relieved for an instant, but this kind of fulfillment of desire leaves only the taste of ashes in one's mouth. The important point is that his behavior tends to be *contrary to interest,* as *senseless* almost as the repetitive beating of one's head against an unyielding stone wall. . . .

Not only are the motives of some mentally ill but noncompulsive wrongdoers *senseless,* they are senseless in the special way that permits us to speak of them as *incoherent.* Their motives do not fit together and make a coherent whole because one kind of desire, conspicuous as a sore thumb, keeps getting in the way. These desires serve ill the rest of their important interests, including their overriding interest in personal integration and internal harmony. They "gum up the works," as we would say of machinery, and throw the person out of "proper working order." The reason they do is that, insofar as these desires are fulfilled, barriers are put in the paths of the others. They are inconsistent with the others in that it is impossible for all to be jointly satisfied, even though it is possible that the others could, in principle, be satisfied together. Moreover, the "senseless" desires, because they do not cohere, are likely to seem alien, not fully expressive of their owner's essential character.[2] When a person acts to satisfy them, it is as if he were acting on somebody else's desires. And, indeed, the alien desires may have a distinct kind of unifying character of their own, as if a new person were grafted on to the old one.

The final and perhaps most important feature common to the examples of voluntary crimes by mentally ill persons is the actor's *lack of insight into his own motive.* The normal person, in rehearsing the possibilities open to him, finds some prospects appealing and others repugnant, and he usually (but not always) knows what it is about a given prospect that makes it appealing or repugnant. If robbing a bank appeals to him, the reason may be that the excitement, the romance, or (far more likely) the money attracts him; and if having more money appeals to him, he usually knows *why* it does too. Normal persons, to be sure, can be mistaken. A criminal may think it is the adventure that is attracting him, instead of the money, or vice versa. It is easy enough to be confused about these things. Often enough we can test our understanding of our own motives by experimental methods. I may think that prospect *X,* which has characteristics *a, b,* and *c,* appeals to me solely because of *a;* but then, to my surprise, prospect *Y,* which has characteristics *a* and *b,* but not *c, repels* me. Hence I conclude that it was not simply the *a*-ness of *X* after all that attracted me. Moreover, even a person who is a model of mental health will be often ignorant or mistaken about the *ultimate* basis of appeal in the things that appeal to him.

The mentally ill person, however, will be radically and fundamentally benighted about the source of the appeal in his immediate objectives, and the truth will be hid from his view by an internal iron curtain. He may think that he is constructed in such a way that little children arouse him sexually, and that is the end of the matter, hardly suspecting that it is the playful, exploratory, irresponsible, and nonthreatening character of his recollected childhood experiences that moves him; or he may think that "exposure for exposure's sake" is what appeals to him in the idea of public undress, whereas really what appeals to him is the public "affirmation of masculinity, a cry of 'Look, here is proof I am a man.' "[3] The true basis of appeal in the criminal's motivation may be, or become, obvious to an outsider, but his illness keeps him blind to it, often, I think, because this blindness is a necessary condition of the appeal itself. At any rate, his lack of self-awareness is no merely contingent thing, like the ignorance that can be charged to absentmindedness, unperceptiveness, objective ambiguities, or the garden varieties of self-deception. The ignorance is the necessary consequence, perhaps even a constituent, of the mental illness, which, taken as a collection of interconnected symptoms, is an alien condition involuntarily suffered.

3

We come back to our original question, then, in a new guise: why should the incoherent and self-concealed character of the mentally ill man's motives be a ground for special consideration when he has voluntarily committed a crime? Perhaps we should enlarge our conception of *compulsion* so that senseless, misunderstood motives automatically count as compulsive. If Jones's chronic desire to do something harmful is as powerful as, but no more powerful than, normal people's desires to do socially acceptable things, then we might think of Jones's desire as a kind of unfair burden. It is no harder for him to restrain on individual occasions, but he must be restraining it *always;* one slip and he is undone. He is really quite unlucky to have this greater burden and danger. The ordinary person is excused when he is made to do what he does not want to do; but the mentally ill man, the argument might go, is excused because of the compulsive weight of his profitless *wants* themselves.

There may be some justice in this argument, but there is little logic. When we being to tamper this profoundly with the concept of compulsion, it is likely to come completely apart. If men can be said to be compelled by their own quite resistible desires, then what is there left to contrast compulsion with?

A more plausible move is to enlarge our conception of what it is to act "in ignorance"—the other category in the Aristotelian formula. The kleptomaniac and the fetishist have no conception of what it is that impels them to their bizarre actions. As we have seen, their conduct may well seem as puzzling to themselves as to any observer. So there is a sense in which they do not know, or realize, what it is they are really doing, and perhaps we should make this ignorance a ground for exculpation; but if we do, we shall be in danger of providing a defense for almost all criminals, normal and ill alike. The bank robber, who is deceived into thinking that it is the adventure that appeals to him when it really is the money, has this excuse available to him, as well as the bully who thinks he inflicts beatings in self-defense when it really is the sight of blood that appeals to him. Lack of insight by itself, then, can hardly be a workable extension of the ignorance defense in courts of law.

It is plain, I think, why the penal law requires rather strict interpretations of compulsion and ignorance. One of its major aims is to deter wrongdoers by providing them with a motive, namely, fear of punishment, which they would not otherwise have for refraining from crime. In close cases involving competent calculators, this new motive might be sufficient to tip the motivational scales toward self-restraint. Mentally ill but rationally competent offenders of the sort I have been discussing, provided only that they *can* restrain themselves, are eminently suited for responsibility because the fear of punishment might make some difference in their behavior. But if they truly cannot help what they do, then the fear of punishment is totally useless and might as well not be induced in them in the first place.

Thus, from the point of view of what punishment can achieve for others, it is a perfectly appropriate mode of treatment for rationally competent, noncompulsive, mentally ill persons. But from the point of view of what can be achieved for the offender himself, I still think it is altogether inappropriate. Some of the aims of an enlightened criminal law,

after all, do concern the offender himself. Sometimes punishment is supposed to "reform" him by intimidation. This no doubt works once in a while for normally prudent and self-interested offenders. For others, greater claims still are made for punishment, which is expected to achieve not merely effective intimidation but also moral regeneration of the offender. But if we treat the mentally ill criminal in precisely the same way as we treat the normal one, we can only bring him to the point of hopeless despair. The prisoner, still devoid of insight into his own motives, will naturally come to wonder how his so-called illness differs from plain wickedness. His bizarre desires will be taken as simply "given," as evil impulses with no point and no reward, simply "there," an integral and irreducible part of himself; and there is no one more pitiably incorrigible than the man convinced of his own intrinsic wickedness and simply resigned to it.

NOTES

1. Alan M. Dershowitz, "The Psychiatrist's Power in Civil Commitment: A Knife that Cuts Both Ways." An abridged version of this talk was published in *Psychology Today,* 2/9 (Feb. 1969), 43–47.
2. Hence the point of the ancient metaphor of "possession."
3. Paul H. Gebhard et al., *Report on Sex Offenders* (New York: Harper & Row, and Paul B. Hoeber, 1965), 399.

REVIEW AND DISCUSSION QUESTIONS

1. Why does mental illness not automatically excuse an agent? What legal view of mental illness do Dershowitz and other writers defend?
2. How do the McNaughten Rules fit with the preceding viewpoint?
3. How do "sick" criminals differ from ordinary ones? What difference is there between mental illness and plain wickedness?
4. Why does Feinberg reject the claim that mental illness is a form of compulsion?
5. According to Feinberg, what are the key features common to voluntary crimes by mentally ill persons? How do these features call into question the idea that insanity is not itself a defense?
6. Is punishment an appropriate response to crimes committed by rationally competent, noncompulsive, mentally ill persons? Explain why or why not.
7. How does Gerber's position in the previous essay compare with Feinberg's position and the position Feinberg criticizes?

Executing Mentally Retarded Murderers

Atkins v. Virginia

The previous two readings discussed mental illness and criminal responsibility. A closely related topic concerns the liability to punishment of mentally retarded defendants. In this case, the U.S. Supreme Court takes up the question of whether the death penalty can be constitutionally imposed on mentally retarded persons. Writing for the majority of the Court, Justice Stevens argues that such executions violate the Eighth Amendment's prohibition of "cruel and unusual" punishments. As proof that use of

536 U.S. 304 (2002).

the death penalty in such cases is out of step with "society's evolving standards of decency," the Court points to recent federal and state laws banning such executions as evidence of an emerging national moral consensus. Although mentally retarded persons can be tried and punished for crimes, they do not have the same moral culpability as most adult criminals. Moreover, their impairments can jeopardize the reliability and fairness of capital proceedings against them. In a vigorous dissent, Justice Scalia denies that there is a national consensus against executing mentally retarded criminals and argues that such executions can be squared with the death penalty's legitimate social purposes.

JUSTICE STEVENS delivered the opinion of the Court: Those mentally retarded persons who meet the law's requirements for criminal responsibility should be tried and punished when they commit crimes. Because of their disabilities in areas of reasoning, judgment, and control of their impulses, however, they do not act with the level of moral culpability that characterizes the most serious adult criminal conduct. Moreover, their impairments can jeopardize the reliability and fairness of capital proceedings against mentally retarded defendants. . . . The American public, legislators, scholars, and judges have deliberated over the question whether the death penalty should ever be imposed on a mentally retarded criminal. The consensus reflected in those deliberations informs our answer to the question presented by this case: whether such executions are "cruel and unusual punishments" prohibited by the Eighth Amendment to the Federal Constitution.

I

Petitioner, Daryl Renard Atkins, was convicted of abduction, armed robbery, and capital murder, and sentenced to death. At approximately midnight on August 16, 1996, Atkins and William Jones, armed with a semiautomatic handgun, abducted Eric Nesbitt, robbed him of the money on his person, drove him to an automated teller machine in his pickup truck where cameras recorded their withdrawal of additional cash, then took him to an isolated location where he was shot eight times and killed.

Jones and Atkins both testified in the guilt phase of Atkins' trial. Each confirmed most of the details in the other's account of the incident, with the important exception that each stated that the other had actually shot and killed Nesbitt. Jones' testimony, which was both more coherent and credible than Atkins', was obviously credited by the jury and was sufficient to establish Atkins' guilt. At the penalty phase of the trial, the State introduced victim impact evidence and proved two aggravating circumstances: future dangerousness and "vileness of the offense." To prove future dangerousness, the State relied on Atkins' prior felony convictions as well as the testimony of four victims of earlier robberies and assaults. To prove the second aggravator, the prosecution relied upon the trial record, including pictures of the deceased's body and the autopsy report.

In the penalty phase, the defense relied on one witness, Dr. Evan Nelson, a forensic psychologist who had evaluated Atkins before trial and concluded that he was "mildly mentally retarded." His conclusion was based on interviews with people who knew Atkins, a review of school and court records, and the administration of a standard intelligence test which indicated that Atkins had a full scale IQ of 59.

The jury sentenced Atkins to death, but the Virginia Supreme Court ordered a second sentencing hearing because the trial court had used a misleading verdict form. At the resentencing, Dr. Nelson again testified. The State presented an expert rebuttal witness, Dr. Stanton Samenow, who expressed the opinion that Atkins was not mentally retarded, but rather was of "average intelligence, at least," and diagnosable as having antisocial personality disorder. The jury again sentenced Atkins to death.

The Supreme Court of Virginia affirmed the imposition of the death penalty. Atkins did not argue before the Virginia Supreme Court that his sentence was disproportionate to penalties imposed for similar crimes in Virginia, but he did contend "that he is mentally retarded and thus cannot be sentenced to death." The majority of the state court rejected this contention. . . . The Court was "not willing to commute Atkins'

sentence of death to life imprisonment merely because of his IQ score." . . .

II

The Eighth Amendment succinctly prohibits "excessive" sanctions. It provides: "Excessive bail shall not be required, nor excessive fines imposed, nor cruel and unusual punishments inflicted." . . . A claim that punishment is excessive is judged not by the standards that prevailed. . . . when the Bill of Rights was adopted, but rather by those that currently prevail. As Chief Justice Warren explained. . . . "The basic concept underlying the Eighth Amendment is nothing less than the dignity of man. . . . The Amendment must draw its meaning from the evolving standards of decency that mark the progress of a maturing society." . . .

III

The parties have not called our attention to any state legislative consideration of the suitability of imposing the death penalty on mentally retarded offenders prior to 1986. In that year, the public reaction to the execution of a mentally retarded murderer in Georgia apparently led to the enactment of the first state statute prohibiting such executions. In 1988, when Congress enacted legislation reinstating the federal death penalty, it expressly provided that a "sentence of death shall not be carried out upon a person who is mentally retarded." In 1989, Maryland enacted a similar prohibition. It was in that year that we decided *Penry,* and concluded that those two state enactments, "even when added to the 14 States that have rejected capital punishment completely, do not provide sufficient evidence at present of a national consensus."

Much has changed since then. Responding to the national attention received by the Bowden execution and our decision in *Penry,* state legislatures across the country began to address the issue. In 1990 Kentucky and Tennessee enacted statutes similar to those in Georgia and Maryland, as did New Mexico in 1991, and Arkansas, Colorado, Washington, Indiana, and Kansas in 1993 and 1994. In 1995, when New York reinstated its death penalty, it emulated the Federal Government by expressly exempting the mentally retarded. Nebraska followed suit in 1998. There appear to have been no similar enactments during the next two years, but in 2000 and 2001 six more States—South Dakota, Arizona, Connecticut, Florida, Missouri, and North Carolina—joined the procession. The Texas Legislature unanimously adopted a similar bill, and bills have passed at least one house in other States, including Virginia and Nevada.

It is not so much the number of these States that is significant, but the consistency of the direction of change. . . . Moreover, even in those States that allow the execution of mentally retarded offenders, the practice is uncommon. Some States, for example New Hampshire and New Jersey, continue to authorize executions, but none have been carried out in decades. Thus there is little need to pursue legislation barring the execution of the mentally retarded in those States. And it appears that even among those States that regularly execute offenders and that have no prohibition with regard to the mentally retarded, only five have executed offenders possessing a known IQ less than 70 since we decided *Penry.* The practice, therefore, has become truly unusual, and it is fair to say that a national consensus has developed against it. . . .

IV

This consensus unquestionably reflects widespread judgment about the relative culpability of mentally retarded offenders, and the relationship between mental retardation and the penological purposes served by the death penalty. Additionally, it suggests that some characteristics of mental retardation undermine the strength of the procedural protections that our capital jurisprudence steadfastly guards.

. . . Mentally retarded persons frequently know the difference between right and wrong and are competent to stand trial. Because of their impairments, however, by definition they have diminished capacities to understand and process information, to communicate, to abstract from mistakes and learn from experience, to engage in logical reasoning, to control impulses, and to understand the reactions

of others. There is no evidence that they are more likely to engage in criminal conduct than others, but there is abundant evidence that they often act on impulse rather than pursuant to a premeditated plan, and that in group settings they are followers rather than leaders. Their deficiencies do not warrant an exemption from criminal sanctions, but they do diminish their personal culpability.

In light of these deficiencies, our death penalty jurisprudence provides two reasons consistent with the legislative consensus that the mentally retarded should be categorically excluded from execution. First, there is a serious question as to whether either justification that we have recognized as a basis for the death penalty applies to mentally retarded offenders. *Gregg* v. *Georgia* (1976) identified "retribution and deterrence of capital crimes by prospective offenders" as the social purposes served by the death penalty. Unless the imposition of the death penalty on a mentally retarded person "measurably contributes to one or both of these goals, it 'is nothing more than the purposeless and needless imposition of pain and suffering,' and hence an unconstitutional punishment.". . .

The reduced capacity of mentally retarded offenders provides a second justification for a categorical rule making such offenders ineligible for the death penalty. The risk "that the death penalty will be imposed in spite of factors which may call for a less severe penalty" is enhanced, not only by the possibility of false confessions, but also by the lesser ability of mentally retarded defendants to make a persuasive showing of mitigation in the face of prosecutorial evidence of one or more aggravating factors. Mentally retarded defendants may be less able to give meaningful assistance to their counsel and are typically poor witnesses, and their demeanor may create an unwarranted impression of lack of remorse for their crimes. . . . Moreover, reliance on mental retardation as a mitigating factor can be a two-edged sword that may enhance the likelihood that the aggravating factor of future dangerousness will be found by the jury. Mentally retarded defendants in the aggregate face a special risk of wrongful execution.

Our independent evaluation of the issue reveals no reason to disagree with the judgment of "the legislatures that have recently addressed the matter" and concluded that death is not a suitable punishment for a mentally retarded criminal. We are not persuaded that the execution of mentally retarded criminals will measurably advance the deterrent or the retributive purpose of the death penalty. Construing and applying the Eighth Amendment in the light of our "evolving standards of decency," we therefore conclude that such punishment is excessive and that the Constitution "places a substantive restriction on the State's power to take the life" of a mentally retarded offender.

The judgment of the Virginia Supreme Court is reversed and the case is remanded for further proceedings not inconsistent with this opinion.

It is so ordered.

Justice Scalia, dissenting: Today's decision is the pinnacle of our Eighth Amendment death-is-different jurisprudence. Not only does it, like all of that jurisprudence, find no support in the text or history of the Eighth Amendment; it does not even have support in current social attitudes regarding the conditions that render an otherwise just death penalty inappropriate. Seldom has an opinion of this Court rested so obviously upon nothing but the personal views of its members.

I

I begin with a brief restatement of facts that are abridged by the Court but important to understanding this case. After spending the day drinking alcohol and smoking marijuana, petitioner Daryl Renard Atkins and a partner in crime drove to a convenience store, intending to rob a customer. Their victim was Eric Nesbitt, an airman from Langley Air Force Base, whom they abducted, drove to a nearby automated teller machine, and forced to withdraw $200. They then drove him to a deserted area, ignoring his pleas to leave him unharmed. According to the co-conspirator, whose testimony the jury evidently credited, Atkins ordered Nesbitt out of the vehicle and, after he had taken only a few steps, shot him one, two, three, four, five, six, seven, eight times in the thorax, chest, abdomen, arms, and legs.

The jury convicted Atkins of capital murder. At resentencing . . . the jury heard extensive evidence of petitioner's alleged mental retardation. . . . The State contested the evidence of retardation. . . .

The jury also heard testimony about petitioner's 16 prior felony convictions for robbery, attempted robbery, abduction, use of a firearm, and maiming. . . . The jury sentenced petitioner to death. The Supreme Court of Virginia affirmed petitioner's sentence.

II

As the foregoing history demonstrates, petitioner's mental retardation was a *central issue* at sentencing. The jury concluded, however, that his alleged retardation was not a compelling reason to exempt him from the death penalty in light of the brutality of his crime and his long demonstrated propensity for violence. . . .

Under our Eighth Amendment jurisprudence, a punishment is "cruel and unusual" if it falls within one of two categories: "those modes or acts of punishment that had been considered cruel and unusual at the time that the Bill of Rights was adopted," and modes of punishment that are inconsistent with modern "standards of decency," as evinced by objective indicia, the most important of which is "legislation enacted by the country's legislatures."

The Court makes no pretense that execution of the mildly mentally retarded would have been considered "cruel and unusual" in 1791. . . . The Court is left to argue, therefore, that execution of the mildly retarded is inconsistent with the "evolving standards of decency that mark the progress of a maturing society." Before today, our opinions consistently emphasized that Eighth Amendment judgments regarding the existence of social "standards" "should be informed by objective factors to the maximum possible extent" and "should not be, or appear to be, merely the subjective views of individual Justices." . . . The Court pays lip service to these precedents as it miraculously extracts a "national consensus" forbidding execution of the mentally retarded from the fact that 18 States—less than *half* (47%) of the 38 States that permit capital punishment (for whom the issue exists)—have very recently enacted legislation barring execution of the mentally retarded. Even that 47% figure is a distorted one. . . . Kansas apparently permits execution of all except the *severely* mentally retarded; New York permits execution of the mentally retarded who commit murder in a correctional facility.

But let us accept, for the sake of argument, the Court's faulty count. That bare number of States alone—*18*—should be enough to convince any reasonable person that no "national consensus" exists. How is it possible that agreement among 47% of the death penalty jurisdictions amounts to "consensus"? Our prior cases have generally required a much higher degree of agreement before finding a punishment cruel and unusual on "evolving standards" grounds. . . .

III

. . . The genuinely operative portion of the opinion, then, is the Court's statement of the reasons why it agrees with the contrived consensus it has found, that the "diminished capacities" of the mentally retarded render the death penalty excessive. . . . The Court gives two reasons why the death penalty is an excessive punishment for all mentally retarded offenders. First, the "diminished capacities" of the mentally retarded raise a "serious question" whether their execution contributes to the "social purposes" of the death penalty, viz., retribution and deterrence. (The Court conveniently ignores a third "social purpose" of the death penalty—"incapacitation of dangerous criminals and the consequent prevention of crimes that they may otherwise commit in the future," *Gregg* v. *Georgia*, . . .).

. . . Surely culpability, and deservedness of the most severe retribution, depends not merely (if at all) upon the mental capacity of the criminal (above the level where he is able to distinguish right from wrong) but also upon the depravity of the crime—which is precisely why this sort of question has traditionally been thought answerable not by a categorical rule of the sort the Court today imposes upon all trials, but rather by the sentencer's weighing of the circumstances (both degree of retardation and depravity of crime) in the particular case. The fact that juries continue to

sentence mentally retarded offenders to death for extreme crimes shows that society's moral outrage sometimes demands execution of retarded offenders. By what principle of law, science, or logic can the Court pronounce that this is wrong? There is none. Once the Court admits (as it does) that mental retardation does not render the offender morally *blameless*, there is no basis for saying that the death penalty is *never* appropriate retribution, no matter *how* heinous the crime. As long as a mentally retarded offender knows "the difference between right and wrong," only the sentencer can assess whether his retardation reduces his culpability enough to exempt him from the death penalty for the particular murder in question.

As for the other social purpose of the death penalty that the Court discusses, deterrence: That is not advanced, the Court tells us, because the mentally retarded are "less likely" than their non-retarded counterparts to "process the information of the possibility of execution as a penalty and . . . control their conduct based upon that information." . . . The Court does not say that *all* mentally retarded individuals cannot "process the information of the possibility of execution as a penalty and . . . control their conduct based upon that information"; it merely asserts that they are "less likely" to be able to do so. But surely the deterrent effect of a penalty is adequately vindicated if it successfully deters many, but not all, of the target class. Virginia's death penalty, for example, does not fail of its deterrent effect simply because *some* criminals are unaware that Virginia *has* the death penalty. . . .

The Court throws one last factor into its grab bag of reasons why execution of the retarded is "excessive" in all cases: Mentally retarded offenders "face a special risk of wrongful execution" because they are less able "to make a persuasive showing of mitigation," "to give meaningful assistance to their counsel," and to be effective witnesses. . . . If this unsupported claim has any substance to it (which I doubt) it might support a due process claim in all criminal prosecutions of the mentally retarded; but it is hard to see how it has anything to do with an *Eighth Amendment* claim that execution of the mentally retarded is cruel and unusual. We have never before held it to be cruel and unusual punishment to impose a sentence in violation of some *other* constitutional imperative. . . .

Today's opinion adds one more to the long list of substantive and procedural requirements impeding imposition of the death penalty imposed under this Court's assumed power to invent a death-is-different jurisprudence. None of those requirements existed when the Eighth Amendment was adopted, and some of them were not even supported by current moral consensus. They include prohibition of the death penalty for "ordinary" murder, for rape of an adult woman, and for felony murder absent a showing that the defendant possessed a sufficiently culpable state of mind; prohibition of the death penalty for any person under the age of 16 at the time of the crime; prohibition of the death penalty as the mandatory punishment for any crime; a requirement that the sentencer not be given unguided discretion, a requirement that the sentencer be empowered to take into account all mitigating circumstances, and a requirement that the accused receive a judicial evaluation of his claim of insanity before the sentence can be executed. There is something to be said for popular abolition of the death penalty; there is nothing to be said for its incremental abolition by this Court.

This newest invention promises to be more effective than any of the others in turning the process of capital trial into a game. One need only read the definitions of mental retardation adopted by the American Association of Mental Retardation and the American Psychiatric Association to realize that the symptoms of this condition can readily be feigned. . . .

I respectfully dissent.

REVIEW AND DISCUSSION QUESTIONS

1. Justice Stevens and the Court majority believe that mentally retarded persons generally know the difference between right and wrong, that they can be competent to stand trial, and that they are not exempt from punishment, but that they nevertheless have diminished culpability. Do you agree or disagree? Explain why.

2. The Court does not appear to distinguish between mild and severe mental retardation. Does the degree of the defendant's retardation matter?
3. If mentally retarded persons are less culpable for their crimes than those who are not retarded, would this rule out capital punishment in all cases? Assuming that we can in general justify executing murderers on the basis of either deterrence or retribution, can either consideration justify executing a murderer who is mentally retarded? If it is wrong to execute people who are mentally retarded, why isn't it also wrong to imprison them?
4. How is your view of this case affected by the fact that the prosecution contested the defense's claim that defendant Atkins is mentally retarded? Will a defendant's mental retardation always be subject to dispute? Is Justice Scalia right to contend that it is easy for a defendant to feign mental retardation?
5. Do you agree that mentally retarded persons accused of murder are less likely to receive the same procedural protections as other defendants and that they face a "special risk of wrongful execution"? If so, does this imply that we should never execute mentally retarded murderers? Could the "special risk" problem be satisfactorily addressed by the courts' taking extra precautions when dealing with mentally retarded defendants?
6. Has a national consensus emerged that it is wrong to execute mentally retarded murderers? Who is right about this—Justice Stevens or Justice Scalia? Does it matter whether there is such a consensus or not?
7. Should Daryl Renard Atkins be executed?

PART FIVE

Philosophical Issues in Civil Law

CHAPTER 16

Compensating for Private Harms: The Law of Torts

This section opens with an essay by Andrew Altman on economic efficiency and the common law. Because it discusses contract and property law as well as tort law, it also serves as a useful introduction to Sections 17 and 18. The remaining essays of this section, however, focus more exclusively on the law of torts and, in particular, the tort of negligence.

A *tort* is a civil wrong or harm, other than a breach of contract, for which the injured party is entitled to compensation. By contrast with the sanctions and penalties of criminal law, tort law provides for redress of the damage one individual suffers at the hands of another. Alongside contract law and property law, tort law is one of the main branches of civil law. It concerns the standards of conduct and rules of liability that govern civil suits between individuals over harms resulting from libel, battery, trespass, nuisance, defective products, automobile accidents, and a wide range of other untoward conduct.

One very important tort is negligence, the main elements of which are explained in the reading by William Prosser. In the famous *Carroll Towing* case, which follows, Judge Leaned Hand presents what has come to be called the "Hand formula" for determining liability in negligence cases. To be liable for negligence, one must, among other things, have exposed others to an unreasonable risk of harm. A plaintiff who has been injured cannot successfully sue a defendant for negligence if the injury resulted from a risk that was too unlikely for the defendant to have anticipated. When is a risk foreseeable? That issue is at the heart of *Stone v. Bolton.* In "Tort Liability and Corrective Justice," Jeffrie G. Murphy and Jules L. Coleman examine some of the philosophical issues that underlie tort liability—in particular, the notion of fault and the requirements of corrective justice—in the context of proposals for no-fault automobile insurance systems. To be liable for negligence, the defendant's conduct must have fallen below the standard of care owed by law to the plaintiff. But when X's negligent conduct toward Y freakishly results in injury to Z, hard questions about tort liability surface. These are taken up in the famous *Palsgraf* case. In the next reading, Stephen R. Perry pursues further the topics of corrective justice and responsibility for our actions, raised in the previous readings. That the defendant caused the plaintiff's injury is, as Prosser explained, a basic element of the tort of

negligence, but what counts as legal or "proximate" cause? In the next reading, Amy Shapiro sheds light on the complicated question of the law's understanding of causation.

This section concludes with several cases that illustrate different aspects of tort law. *Summers v. Tice* is well known because it permitted a defendant to be sued for negligence even though his actions could not be shown to have caused the harm suffered by the plaintiff. *Yania v. Bigan* and its companion cases address the issue of whether and when one has a duty to rescue another. If a defendant is to be held liable for negligence, he or she must have harmed the plaintiff in some way. Can life itself be a harm? The final case, *Berman v. Allen*, takes up this issue.

The Efficiency of the Common Law

Andrew Altman

In this essay, Andrew Altman illustrates how economic principles, in particular the goal of economic efficiency, apply to and illuminate central elements of contract law, property law, and tort law (in particular, negligence), thus providing a useful introduction to the three subject areas of Part V of this book. In Part III, Section 11, we saw that the "law and economics" approach to jurisprudence is controversial and that it may not supply a complete account of the nature of law, but there is no doubt that looking at the economic aspect of contract, tort, and property law sheds light on some of their key features and central doctrines. (The reader may wish to consult Altman's essay in Section 11 for further exposition and critical analysis of the law and economics approach.)

The common law consists of the judge-made system of legal rules that govern property, contracts, and torts. Although much of it was originally inherited from England, the common law underwent modification and development at the hands of American judges. Economic analysts of law often claim that the rules of common law, on the whole, lead to an economically efficient allocation of resources.

CONTRACT LAW

Promoting Efficient Exchange

In order to promote efficiency, a system of legal rules must ensure that resources end up in the hands of those who place the most value on them. If you value my wheat more than I do and I value your corn more than you do, it is inefficient if I keep the wheat and you keep the corn. That is because the two resources, wheat and corn, are not in the hands of the people who value them the most. Economic analysts argue that market transactions are vital in promoting efficiency because the transactions help ensure that resources end up in the hands of those who value them the most.

Accordingly, an efficient system of law must have rules that enable persons voluntarily to exchange the resources they possess. In the common-law system, according to economic analysis, the rules of contract law help to ensure that resources get transferred to those who value them the most. The rules are especially important in those situations where the parties to a contract do not simultaneously exchange their respective goods.

For example, suppose that my wheat is ready for harvesting and delivery now but that your corn will not be ready until next month. We might enter a contract in which I deliver the wheat to you before you deliver the corn to me. When I deliver the wheat, though, I am in a vulnerable situation. Now that you have my wheat, you might be tempted to hold onto your corn. The rules of contract law aim to deter such inefficient behavior on your part and thus give me sufficient assurance to send you my wheat. Only in that way will the economically efficient transaction occur. Accordingly, the rules of contract law provide

that you must pay damages to me in case you breach our contract.

Suppose, however, that someone else who values your corn even more than I do gets in contact with you after I have delivered my wheat to you but before you have delivered your corn to me. She makes you a better offer for the corn than I made. Thus, it would be more economically efficient for that person to get the corn than for me to do so: she values it more. Does the law allow that to happen? Yes, because contract law only requires you to pay me the cash value of the corn to me. If you get a better offer for the corn from someone else, then you can pay me that cash amount and still have some left over for yourself. Everyone ends up better off, and economic efficiency is promoted.

Certain kinds of contracts are legally unenforceable, for example, those made as a result of duress. If you coerce me into signing a contract with you at gunpoint, the law will refuse to enforce the contract. Economic analysts point out that the rule regarding duress is economically efficient. If I do not willingly enter into a contract with you for something I possess, economic analysis will conclude that you do not really value the item more than I do. If you did, then you would voluntarily pay me for it. For the economic analyst, a person's willingness to pay is the most reliable evidence of the value of something to that person.

According to economic analysis, it is not efficient for society's legal system to enforce every promise or agreement, even when we ignore the ones made under duress. Enforcement uses society's resources, and if it costs more to enforce a contract or promise than the net gain that accrues as a result of meeting its terms, society will have lost more than it gains. Economic analysis claims that contract law contains a doctrine that helps society avoid the inefficient enforcement of such promises and contracts, the doctrine of consideration.

THE DOCTRINE OF CONSIDERATION

Consideration is defined by the law as something of value that is given or promised in exchange for something of value. If I offer to sell you my car and you promise to pay me $500, then your $500 and my car both count as consideration: in consideration of your money, I will give you my car, and in consideration of my car, you will give me your money. Consideration is the essence of the deal we make.

The doctrine of consideration says that in order for a promise or agreement to be legally binding, each party must give (or promise) consideration for what the other gives (or promises). This means that "gratuitous" promises are not legally binding. Thus, suppose I promise to give you $10 so you can buy a new compact disc, and I ask for nothing (and you offer nothing) in return. The promise is without consideration and so is gratuitous. It is not legally binding and so, in general, I may change my mind the next day without incurring legal liability.

Does the doctrine of consideration promote efficiency? Economic analysts say yes. They claim that it excludes from legal enforcement relatively trivial "social" promises whose enforcement would be economically inefficient. After all, it uses up society's resources when you take me to court in a dispute over my promise to give you $10 to buy a compact disc or when Joe takes Jane to court in order to settle a dispute about whether she promised to go out on a date with him. The valuable time of judges, lawyers, court employees could be more efficiently spent if such informal social promises were denied legal enforcement, which is what the doctrine of consideration does.

Economic analysts also argue that the doctrine excludes from enforcement those agreements that are so vague that the courts would, in effect, be required to define the terms. In order for there to be consideration, there must be a definite something—a sum of money or a specific item or a specific promise of such—that is given in the exchange. Where nothing definite is given or promised, the courts have no specific bargain to enforce. That means that the courts would have to set the terms of the bargain in order to enforce it. But that is largely inefficient because the parties themselves are in a much better position than the courts to say what contract terms make them better off. The parties know better what value they place on things and what it is worth to them to obtain the item that the other party is offering. As a general rule, efficiency demands that the parties to a contract set the terms themselves and that courts restrict themselves to enforcing the terms that the parties have set. Accordingly, the doctrine of consideration helps promote efficiency by telling potential contract partners, "You need to settle on the specific terms of

the contract yourselves, and, if you fail to do so, we are not going to formulate the terms for you."

Contract law generally requires that there be consideration for a promise to be legally enforceable. But it also keeps courts out of the business of evaluating whether the consideration given for a particular promise is adequate. The law is concerned with the existence of consideration but not with its adequacy. In other words, the courts will not inquire into whether each of the parties to a contract is getting his or her money's worth.

Economic analysts argue that the law's refusal to consider adequacy is efficient. This is due to the fundamental economic principle that individuals are in the best position to say what something is worth to them. The principle helps explain why it is efficient for courts to insist on definite consideration before they will enforce a contract, but it also helps explain why it is efficient for them to avoid questions about the adequacy of consideration. If you are the one in the best position to say what something offered as part of a contract is worth to you, then the courts should not second-guess whether you are getting your money's worth.

NEGLIGENCE AND THE HAND FORMULA

The common law not only sets the rules for the enforceable bargains people make. One of its important functions is to enable those injured by another's negligence to gain compensation. The tort law sets the rules covering negligent behavior, including the criteria that determine what behavior counts as negligent.

Negligence is ordinarily explained as the failure to exercise reasonable care in one's activities. But what level of care is reasonable? Economic analysts claim that the idea of efficiency can help answer that question. Reasonable care, they say, is economically efficient care.

Suppose that unusually cold weather causes the underground water pipe that I installed to burst, causing damage to your property. Was I negligent in not putting the pipe deeper in the ground, where it would not have frozen and burst? The economic analyst will answer this question by comparing the cost of putting the pipe deeper to the *expected* cost of the burst pipe. The expected cost of the burst pipe is obtained by multiplying the probability that the pipe would burst by the harm done. Suppose that the probability of such a burst was 0.01 and that $1,000 worth of harm was caused. That means that the expected cost of the harm was $10. Suppose that it would have cost me an extra $20 to have put the pipe deep enough to avoid the bursting. Then it would *not* have been "cost justified" to have placed the pipe that deep: it would have cost more than its expected benefit of saving $10. Taking precautions that are not cost justified is not economically efficient.

Suppose, however, that it would have only cost me $8 to have placed the pipe deep enough to avoid the bursting. Then it *would* have been cost justified to place the pipe that deep, and for that reason it would have been economically efficient to do so. Thus, I was acting in an inefficient way by not placing the pipe that deep, and according to the economic interpretation of negligence, I was being negligent.

The economic interpretation of negligence was first proposed by the famous judge Learned Hand. Judge Hand ruled that a defendant was negligent if the cost of precautions that could have prevented the harm she caused was less than the expected harm. Thus, the so-called "Hand formula" for negligence is that a defendant is negligent if and only if B is less than PI, where B = the burden of preventative measures, P = the probability of harm, and I = the injury caused. . . .

On the economic view, then, the common-law doctrine of negligence promotes efficiency by encouraging persons to take cost-justified measures to prevent harm to others. It holds them liable for carelessness if but only if they failed to invest resources in prevention up to the point where additional investment would cost as much or more than the expected marginal savings.

PROPERTY

In addition to negligence and contracts, an important part of the common law covers the ownership and use of property. Economic analysts claim that these common-law rules are efficient as well.

Consider the rule that a property owner owes no duty of care to a trespasser. If I own land that has hidden pits and crevices on it, the common law

does not require me to post signs that would warn trespassers of the danger. Suppose you come onto my land without my consent and fall into a pit, injuring yourself. Under the common law, I am not liable for the harm, since you are a trespasser, regardless of whether I was negligent in failing to fill in the pit or take other precautionary measures.

The economic understanding for the rule is that it promotes efficiency because the trespasser can generally avoid the cost of injury more cheaply than the landowner can: he need only refrain from trespassing, which will usually cost less than my filling the pit, putting a fence around it or taking some similar precautionary measure. Requiring the trespasser to shoulder the costs of the injury creates an incentive for potential trespassers to refrain from that kind of illegal action. And since the potential trespasser can take measures to avoid the costs of trespassing injuries more cheaply than can landowners, efficiency is thereby promoted, according to economic analysts.

The trespassing example can be generalized to cases in which both parties are engaged in perfectly legitimate activities. In many types of disputes, the uses to which two people put their respective pieces of property end up causing harm to one of the properties. It is efficient for the harm to be avoided, but this can be accomplished only if one of the parties alters his use of his property. Economic analysis says that in such cases of "incompatible" property use, the common law typically assigns rights to the parties based on which one can take steps that avoid the harm more cheaply. The cheaper "cost avoider" is given the burden of altering his use of his property, while the other party is given the legal right to continue using his property as he has been.

Suppose a railroad emits sparks and that, once in a while, the sparks will cause a fire that burns down the crops on a farm adjacent to the railroad. The railroad could prevent the sparks by slowing its speed by 15 mph. But that slowdown will cost the railroad $100 in revenues. The farmer could prevent the fires by not planting on land within thirty feet of the tracks. But not planting on that land will cost the farmer $50 in revenues. The common law will, according to economic analysis, give the railroad the right to emit sparks and put the burden of avoiding fires on the farmer, the cheaper cost avoider. The farmer will have no choice but to refrain from planting crops within thirty feet of the tracks.

procedure goes to the wealthy

REVIEW AND DISCUSSION QUESTIONS

1. How does contract law promote efficient exchange? Why are some contracts legally unenforceable?
2. What is the doctrine of consideration, and how does it promote economic efficiency? Why, according to economic analysts, does the law refuse to consider the adequacy of consideration?
3. How do the economic approach and, in particular, the "Hand formula" illuminate the legal concept of negligence?
4. In property law, how do the trespass rule and the handling of "incompatible" property use promote economic efficiency?
5. Are there ways in which contract, tort, and property law fail to promote economic efficiency? Is this the only value that the law in fact promotes? That it should promote?

Negligence

William Prosser

Torts are private legal actions claiming compensation for damages. Some torts concern intentional injuries—for example, punching someone in the face or spreading lies about the person. Negligence as a tort, however, concerns unintentional injuries—for example, accidentally running over a sunbather with your dune buggy. Although it was not distinguished as a separate and distinct ground of legal

liability until the early nineteenth century, today negligence is probably the most important and certainly the most frequently litigated category of torts. In the following essay, William Prosser, one of the foremost authorities on the law of torts, spells out the main elements of the tort of negligence, that is, the circumstances under which one can be held legally liable for negligent conduct. He explains what the law means by negligence and explicates such crucial concepts as that of unreasonable risk and the standard of the "reasonable man."

About the year 1825, negligence began to be recognized as a separate and independent basis of tort liability. Its rise coincided in a marked degree with the Industrial Revolution; and it very probably was stimulated by the rapid increase in the number of accidents caused by industrial machinery, and in particular by the invention of railways. . . .

Intentional injuries, whether direct or indirect, began to be grouped as a distinct field of liability, and negligence remained as the main basis for unintended torts. Today it is not at all disputed that separate problems and principles, as well as distinct questions of policy, arise in negligence cases.

ELEMENTS OF CAUSE OF ACTION

Negligence, as we shall see, is simply one kind of conduct. But a cause of action founded upon negligence, from which liability will follow, requires more than conduct. The traditional formula for the elements necessary to such a cause of action may be stated briefly as follows:

1. A duty, or obligation, recognized by the law, requiring the actor to conform to a certain standard of conduct, for the protection of others against unreasonable risks.
2. A failure on his part to conform to the standard required. These two elements go to make up what the courts usually have called negligence; but the term quite frequently is applied to the second alone. Thus it may be said that the defendant was negligent, but is not liable because he was under no duty to the plaintiff not to be.
3. A reasonably close causal connection between the conduct and the resulting injury. This is what is commonly known as "legal cause," or "proximate cause."
4. Actual loss or damage resulting to the interests of another. Since the action for negligence developed chiefly out of the old form of action on the case, it retained the rule of that action, that proof of damage was an essential part of the plaintiff's case. Nominal damages, to vindicate a technical right, cannot be recovered in a negligence action, where no actual loss has occurred. The threat of future harm, not yet realized, is not enough. Negligent conduct in itself is not such an interference with the interests of the world at large that there is any right to complain of it, or to be free from it, except in the case of some individual whose interests have suffered.

Such a statement must, however, be qualified to the extent that, as in the case of other torts, where irreparable injury is threatened, a court of equity may act by injunction to prevent the harm before it occurs. Even here the damage, even though only potential, is the basis for granting relief. . . .

Negligence is a matter of risk—that is to say, of recognizable danger of injury. It has been defined as "conduct which involves an unreasonably great risk of causing damage," or, more fully, conduct "which falls below the standard established by law for the protection of others against unreasonably great risk of harm." "Negligence is conduct, and not a state of mind." In most instances, it is caused by heedlessness or carelessness, which makes the negligent party unaware of the results which may follow from his act. But it may also exist where he has considered the possible consequences carefully, and has exercised his own best judgment. The standard imposed by society is an external one, which is not necessarily based upon any moral fault of the individual; and a failure to conform to it is negligence, even though it may be due to stupidity, forgetfulness, an excitable temperament, or even sheer ignorance. The almost universal use of the phrase "due care" to describe conduct which is not negligent, should not be permitted to obscure the fact that the real basis of negligence is not carelessness, but behavior which should be recognized as involving unreasonable danger to others.

Previous reference has been made to the distinction between negligence and intent. In negligence, the actor does not desire to bring about the consequences which follow, nor does he know that they are substantially certain to occur, or believe that they will. There is merely a risk of such consequences, sufficiently great to lead a reasonable man in his position to anticipate them, and to guard against them. If an automobile driver runs down a man in the street before him, with the desire to hit him, or with the belief that he is certain to do so, it is an intentional battery; but if he has no such desire or belief, but merely acts unreasonably in failing to guard against a risk which he should appreciate, it is negligence. As the probability of injury to another, apparent from the facts within his knowledge, becomes greater, his conduct takes on more of the attributes of intent, until it reaches that substantial certainty of harm which juries, and sometimes courts, may find inseparable from intent itself. Such intermediate mental states, based upon a recognizable great probability of harm, may still properly be classed as "negligence," but are commonly called "reckless," "wanton," or even "willful." They are dealt with, in many respects, as if the harm were intended, so that they become in effect a hybrid between intent and negligence, occupying a sort of penumbra between the two. They will be dealt with in a later section.

Negligence already has been defined as conduct which falls below a standard established by the law for the protection of others against unreasonable risk of harm. The idea of risk necessarily involves a recognizable danger, based upon some knowledge of the existing facts, and some reasonable belief that harm may follow. A risk is a danger which is apparent, or should be apparent, to one in the position of the actor. The culpability of the actor's conduct must be judged in the light of the possibilities apparent to him at the time, and not by looking backward "with the wisdom born of the event." The standard must be one of conduct, rather than of consequences. It is not enough that everyone can see now that the risk was great, if it was not apparent when the conduct occurred. The court must put itself in the actor's place. At the same time, the standard imposed must be an external one, based upon what society demands of the individual, rather than upon his own notions of what is proper. An honest blunder, or a mistaken belief that no damage will result, may absolve him from moral blame, but the harm to others is still as great, and the actor's individual standards must give way to those of the public. In other words, society may require of him not to be a fool.

In the light of the recognizable risk, the conduct, to be negligent, must be unreasonable. Nearly all human acts, of course, carry some recognizable but remote possibility of harm to another. No man so much as rides a horse without some chance of a runaway, or drives a car without the risk of a broken steering gear or a heart attack. But these are not unreasonable risks. Those against which the actor is required to take precautions are those which society, in general, considers sufficiently great to demand them. No man can be expected to guard against harm from events which are not reasonably to be anticipated at all, or are so unlikely to occur that the risk, although recognizable, would commonly be disregarded. An unprecedented frost or flood, an automobile thrown unexpectedly against a pillar on the corner, a child picking up a plank with a nail in it and dropping it on his foot, a pedestrian slipping on a small bit of gravel in the highway, the ricochet of a bullet at an almost impossible angle—all of these things have happened, and will occur again; but they are not so likely to do so on any particular occasion as to make it necessary to burden the freedom of human action with precautions against them. Such events are regarded as "unavoidable accidents," for which there is no liability.

On the other hand, if the risk is an appreciable one, and the possible consequences are serious, the question is not one of mathematical probability alone. The odds may be a thousand to one that no train will arrive at the very moment that an automobile is crossing a railway track, but the risk of death is nevertheless sufficiently serious to require the driver to look for the train. It may be highly improbable that lightning will strike at any given place or time; but the possibility is there, and it requires precautions for the protection of inflammables. As the gravity of the possible harm increases, the apparent likelihood of its occurrence need be correspondingly less.

Against this probability, and gravity, of the risk, must be balanced in every case the utility of the type of conduct in question. The problem is whether "the game is worth the candle." Many risks

may reasonably be run, with the full approval of the community. Chief among the factors which must be considered is the social value of the interest which the actor is seeking to advance. A man may be justified in dashing into the path of a train to save the life of a child, where it would be arrant folly to save his hat. A railway will be permitted, or even required, to blow a whistle to warn travelers at a crossing, although it is likely to frighten horses on the highway; it may be negligence to blow the same whistle without the same occasion for warning. The public interest will justify the use of dangerous machinery, so long as the benefits out-weigh the risk, and a railroad may reasonably be constructed near a highway, even at the expense of some danger to those who use it. . . .

Consideration must also be given to any alternative course open to the actor. Whether it is reasonable to travel a dangerous road may depend upon the disadvantages of another route; and while mere inconvenience or cost may not in themselves be sufficient to justify proceeding in the face of great danger, they may justify taking other risks which are not too extreme. A county will not be required, at ruinous expense, to build a bridge which will be safe against any accident that might be anticipated; but the converse is also true, and where it can cheaply and easily post a warning, it may be required to do so. A railroad need not do without a turntable because there is some chance that children will play on it and be hurt; but it is quite another matter to keep it locked.

The alternative dangers to the actor himself and to others must be thrown into the scale, and a balance struck in which all of these elements are weighed.

It is fundamental that the standard of conduct which is the basis of the law of negligence is determined by balancing the risk, in the light of the social value of the interest threatened, and the probability and extent of the harm, against the value of the interest which the actor is seeking to protect, and the expedience of the course pursued. For this reason, it is seldom possible to reduce negligence to any definite rules; it is "relative to the need and the occasion," and conduct which would be proper under some circumstances becomes negligence under others.

THE REASONABLE MAN

The whole theory of negligence presupposes some uniform standard of behavior. Yet the infinite variety of situations which may arise makes it impossible to fix definite rules in advance for all conceivable human conduct. The utmost that can be done is to devise something in the nature of a formula, the application of which in each particular case must be left to the jury, or to the court. The standard of conduct which the community demands must be an external and objective one, rather than the individual judgment, good or bad, of the particular actor; and it must be, so far as possible, the same for all persons, since the law can have no favorites. At the same time, it must make proper allowance for the risk apparent to the actor, for his capacity to meet it, and for the circumstances under which he must act.

The courts have dealt with this very difficult problem by creating a fictitious person, who never has existed on land or sea: the "reasonable man of ordinary prudence." Sometimes he is described as a reasonable man, or a prudent man, or a man of average prudence, or a man of ordinary sense using ordinary care and skill. It is evident that all such phrases are intended to mean very much the same thing. The actor is required to do what such an ideal individual would be supposed to do in his place. A model of all proper qualities, with only those human shortcomings and weaknesses which the community will tolerate on the occasion, "this excellent but odious character stands like a monument in our Courts of Justice, vainly appealing to his fellow-citizens to order their lives after his own example."

The courts have gone to unusual pains to emphasize the abstract and hypothetical character of this mythical person. He is not to be identified with any ordinary individual, who might occasionally do unreasonable things; he is a prudent and careful man, who is always up to standard. Nor is it proper to identify him even with any member of the very jury who are to apply the standard; he is rather a personification of a community ideal of reasonable behavior, determined by the jury's social judgment. . . .

Physical Attributes

The conduct of the reasonable man will vary with the situation with which he is confronted. The jury must

therefore be instructed to take the circumstances into account; negligence is a failure to do what the reasonable man would do "under the same or similar circumstances." Under the latitude of this phrase, the courts have made allowance not only for the external facts, but for many of the characteristics of the actor himself, and have applied, in many respects, a more or less subjective standard. "It would appear that there is no standardized man; that there is only in part an objective test; that there is no such thing as reasonable or unreasonable conduct except as viewed with reference to certain qualities of the actor—his physical attributes, his intellectual powers, probably, if superior, his knowledge and the knowledge he would have acquired had he exercised standard moral and at least average mental qualities at the time of action or at some connected time."

As to his physical characteristics, the reasonable man may be said to be identical with the actor. The man who is blind or deaf, or lame, or is otherwise physically disabled, is entitled to live in the world and to have allowance made by others for his disability, and he cannot be required to do the impossible by conforming to physical standards which he cannot meet. Similar allowance has been made for the weaknesses of age and sex. At the same time, the conduct of the handicapped individual must be reasonable in the light of his knowledge of his infirmity, which is treated merely as one of the circumstances under which he acts. A blind man may be negligent in going into a place of known danger, just as one who knows that he is subject to epileptic fits, or is about to fall asleep, may be negligent in driving a car. It is sometimes said that a blind man must use a greater degree of care than one who can see; but it is now generally agreed that as a fixed rule this is inaccurate, and that the correct statement is merely that he must take the precautions, be they more or less, which the ordinary reasonable man would take if he were blind. In theory the standard remains the same, but it is sufficiently flexible to take his physical defects into account.

Mental Capacity

As to the mental attributes of the actor, the standard remains of necessity an external one. "The law," says Mr. Justice Holmes in a much quoted passage, "takes no account of the infinite varieties of temperament, intellect, and education which make the internal character of a given act so different in different men. It does not attempt to see men as God sees them, for more than one sufficient reason." The fact that the individual is a congenital fool, cursed with inbuilt bad judgment, or that in the particular instance he "did not stop to think," or that he is merely a stupid ox, or of an excitable temperament which causes him to lose his head and get "rattled," obviously cannot be allowed to protect him from liability. Apart from the very obvious difficulties of proof as to what went on in his head, it may be no bad policy to hold a fool according to his folly. The harm to his neighbors is quite as great, and may be greater, than if he had a modicum of brains; and if he is to live in the community, he must learn to conform to its standards or pay for what he breaks. As to all such mental deficiency, no allowance is made; the standard of reasonable conduct is applied, and "it is not enough that the defendant did the best he knew how."

Obviously, however, an extreme is reached at which the mental deficiency has prevented the individual from comprehending the danger, or from taking action to avoid it; and the question then becomes, how can we find the tort of negligence at all? Can a complete imbecile, for example, be negligent? It is a rather mysterious fact that no case can be found involving such a condition on the part of a defendant. There are, however, quite a few cases dealing with contributory negligence of plaintiffs; and in all of them except one, it has been held that the jury could find that there was no negligence. . . .

In cases in which the mental aberration has reached the point of actual insanity, the tendency has been to carry over the standard of the reasonable man, and to apply it even to the negligence of those who are definitely insane. This has not gone without criticism, and there has been much argument that the exception should be made when mental deficiency reaches such an extreme. There are a handful of contributory negligence cases in which the mental state has prevented the individual from comprehending or avoiding the danger, and the jury has been permitted to find that there is no negligence. There are two

decisions in Wisconsin and Ontario which have come to the same conclusion as to insane defendants. It seems rather probable that the cases as to lunatics will follow the line of mental deficiency in general.

On the other hand a transitory unconsciousness, or delirium due to illness, commonly is regarded as a "circumstance" depriving the actor of control over his conduct, which will relieve him of liability.

Whether intoxication is to be regarded as a physical or a mental disability is probably of no importance at all to anyone. On either basis, it is common enough; and it is uniformly held that voluntary or negligent intoxication cannot serve as excuse or absolution for acts done in that condition which would otherwise be negligent. One good reason is that such an excuse would be far too common and too easy to assert; another is that drunkenness is so anti-social that one who indulges in it ought to be held to the consequences. It is sometimes said that such intoxication is negligence in itself; but this is scarcely correct, since a drunken man may still behave in a perfectly reasonable manner. The proper statement would seem to be that one who intentionally or negligently becomes intoxicated is held thereafter to the same standard of conduct as if he were sober. . . .

Knowledge

One of the most difficult questions in connection with negligence is that of what the actor may be required to know. . . . So far as perception is concerned, it seems clear that, unless his attention is legitimately distracted, the actor must give to his surroundings the attention which a standard reasonable man would consider necessary under the circumstances, and that he must use such senses as he has to discover what is readily apparent. He may be negligent in failing to look, or in failing to observe what is visible when he does look. As to memory, he is required to fix in his mind those matters which would make such an impression upon the standard man, and, unless he is startled, or his attention is distracted for some sufficient reason, to bear them in mind, at least for a reasonable length of time.

The real difficulty lies with the question of experience. The late Henry T. Terry came to the conclusion that "there are no facts whatever which every person in the community is absolutely bound at his peril to know." It seems clear, however, that there are certain things which every adult with a minimum of intelligence must necessarily have learned: the law of gravity, the fact that fire burns and water will drown, that inflammable objects will catch fire, that a loose board will tip when it is trod on, the ordinary features of the weather to which he is accustomed, and similar phenomena of nature. He must know in addition a few elementary facts about himself: the amount of space he occupies, the principles of balance and leverage as applied to his own body, and, to the extent that it is reasonable to demand it of him, the limits of his own strength, as well as some elementary rules of health.

But beyond this, it seems clear that any individual who has led a normal existence will have learned much more: the traits of common animals, the normal habits, capacities and reactions of other human beings, including their propensities toward negligence and crime, the danger involved in explosives, inflammable liquids, electricity, moving machinery, slippery surfaces and firearms, the fact that an automobile is not easy to control in deep sand, that worn tires will blow out, and many other things. Such an individual will not be credited or excused when he denies knowledge of the risk; and to this extent, at least, there is a minimum standard of knowledge, based upon what is common to the community.

The few cases which have considered the question have held that when an abnormal individual who lacks the experience common to the particular community comes into it, as in the case of the old lady from the city who comes to the farm without ever having learned that a bull is a dangerous beast, the standard of knowledge will still be applied, and it is the individual who must conform to the community, rather than vice versa.

Above this minimum, once it is determined, the individual will not be held to knowledge of risks which are not known or apparent to him. He may, however, know enough to be conscious of his own ignorance, and of possible danger into which it may lead him; and if that is the case, as where a layman attempts to give medical treatment, or one enters a strange dark passage, or an

automobile driver proceeds with a mysterious wobble in his front wheels, or traverses a strange town without an attempt to discover the meaning of unfamiliar purple traffic lights which suddenly confront him, he will be found negligent in proceeding in the face of known ignorance. He may, furthermore, be engaged in an activity, or stand in a relation to others, which imposes upon him an obligation to investigate and find out, so that he becomes liable not so much for being ignorant as for remaining ignorant; and this obligation may require him to know at least enough to conduct an intelligent inquiry as to what he does not know. The occupier of premises who invites business visitors to enter, the landlord who installs a gas heater in a bathroom used by his tenants, the telephone company which erects wires in the street, the manufacturer of goods to be sold to the public, the carrier who undertakes to transport passengers, all are charged with the duty of the affirmative action which would be taken by a reasonable man in their position, to discover dangers of which they may not be informed. As scientific knowledge advances, and more and more effective tests become available, what was excusable ignorance yesterday becomes negligent ignorance today.

Superior Knowledge, Skill, and Intelligence

Thus far the question has been one of a minimum standard, below which the individual will not be permitted to fall. But if he has in fact knowledge, skill, or even intelligence superior to that of the ordinary man, the law will demand of him conduct consistent with it. The vendor of fur coats who has learned from experience that some few persons are especially susceptible to dermatitis caused by a particular dye must take precautions which might not be required if he had remained in ignorance. Upon the same basis, a physician who is possessed of unusual skill or knowledge must use care which is reasonable in the light of his special ability and information, and may be negligent where an ordinary doctor would not.

Professional men in general, and those who undertake any work calling for special skill, are required not only to exercise reasonable care in what they do, but also to possess a standard minimum of special knowledge and ability. Most of the decided cases have dealt with physicians and surgeons, but the same is undoubtedly true of dentists, pharmacists, psychiatrists, attorneys, architects and engineers, accountants, abstracters of title, and many other professions and even skilled trades. Since, allowing for the inevitable differences in the work done, the principles applied to all of these appear to be quite identical, and since the medical cases are by far the most numerous, it will be convenient to talk only of physicians and surgeons.

A physician may, although he seldom does, contract to cure his patient, or to accomplish a particular result, in which case he may be liable for breach of contract when he does not succeed. In the absence of such an express agreement, he does not warrant or insure the outcome of his treatment, and he will not be liable for an honest mistake of judgement, where the proper course is open to reasonable doubt. But by undertaking to render medical services, even though gratuitously, he will ordinarily be understood to hold himself out as having standard professional skill and knowledge. The formula under which this usually is put to the jury is that he must have the skill and learning commonly possessed by members of the profession in good standing; and he will be liable if harm results because he does not have them. Sometimes this is called the skill of the "average" member of the profession; but this is clearly misleading, since only those in good professional standing are to be considered; and of these it is not the middle but the minimum common skill which is to be looked to. If the defendant represents himself as having greater skill than this, as for example where he holds himself out as a specialist, or as having less, and the patient accepts treatment with that understanding, the standard is modified accordingly.

The courts have been compelled to recognize that there are areas in which even experts will disagree. Where there are different schools of medical thought, it is held that the dispute cannot be settled by the law, and the doctor is entitled to be judged according to the tenets of the school he professes to follow. This does not mean, however, that any quack, charlatan or crackpot can set himself up as a "school," and so apply his individual

ideas without liability. A "school" must be a recognized one with definite principles, and it must be the line of thought of at least a respectable minority of the profession. In addition, there are minimum requirements of skill and knowledge as to both diagnosis and treatment, particularly in the light of modern licensing statutes, which anyone who holds himself out as competent to treat human ailments is required to have, regardless of his personal views on medical subjects. Furthermore the physician is required to exercise reasonable care in ascertaining the operational facts upon which his diagnosis is based, and will be liable if he fails to do so.

REVIEW AND DISCUSSION QUESTIONS

1. What four basic elements are necessary to establish negligence as a legal cause of action? Give examples of conduct that is not negligent because one of these elements is missing.
2. Why is it difficult to reduce negligence to a specific set of rules?
3. Who or what is the "reasonable man"? How do the courts use the concept? What physical and mental characteristics does the "reasonable man" have? How does the reasonable-man standard apply to the conduct of a blind or handicapped person? A stupid or hotheaded person? What difficulties or limitations, if any, do you see in the standard?
4. How does the law treat those such as physicians who have superior knowledge or skill?
5. Do you see any ways in which the law's treatment of negligence differs from common sense or our ordinary ideas of fault and responsibility?

Economic Efficiency and the "Hand Formula"

United States v. Carroll Towing Company

A tugboat owned by Carroll Towing was moving a line of barges when one barge (the *Anna C*) owned by Conners Company broke away and crashed into another ship. The *Anna C* leaked oil, and the U.S. government, the owner of the oil, sued Carroll Towing (owner of the tug who pulled *Anna C*) for the oil. The issue in this case is whether the *Anna C*'s owners (Connors Co.) were partly negligent for the lost oil in the barge because they did not have anybody on board who could prevent the *Anna C* from crashing into another ship. In this opinion, Judge Learned Hand propounds what has since come to be called the "Hand formula" (discussed by Altman earlier).

L. Hand, C. J.: It appears from the foregoing review that there is no general rule to determine when the absence of a bargee or other attendant will make the owner of the barge liable for injuries to other vessels if she breaks away from her moorings. However, in any cases where he would be so liable for injuries to others, obviously he must reduce his damages proportionately, if the injury is to his own barge. It becomes apparent why there can be no such general rule, when we consider the grounds for such a liability. Since there are occasions when every vessel will break from her moorings, and since, if she does, she becomes a menace to those about her, the owner's duty, as in other similar situations, to provide against resulting injuries is a function of three variables: (1) the probability that she will break away; (2) the gravity of the

159 F.2d 169 (2d Cir. 1947).

resulting injury, if she does; (3) the burden of adequate precautions. Possibly it serves to bring this notion into relief to state it in algebraic terms: if the probability be called *P;* the injury, *I;* and the burden, *B;* liability depends upon whether *B* is less than *I*, multiplied by *P:* i.e., whether $B < PI$ Applied to the situation at bar, the likelihood that a barge will break from her fasts and the damage she will do, vary with the place and time; for example, if a storm threatens, the danger is greater; so it is, if she is in a crowded harbor where moored barges are constantly being shifted about. On the other hand, the barge must not be the bargee's prison, even though he lives aboard; he must go ashore at times. We need not say whether, even in such crowded waters as New York Harbor a bargee must be aboard at night at all; it may be that the custom is otherwise, as Ward, J., supposed in *The Kathryn B. Guinan*, 176 F.2d 301; and that, if so, the situation is one where custom should control. We leave that question open; but we hold that it is not in all cases a sufficient answer to a bargee's absence without excuse, during working hours, that he has properly made fast his barge to a pier, when he leaves her. In the case at bar the bargee left at five o'clock in the afternoon of January 3rd, and the flotilla broke away at about two o'clock in the afternoon of the following day, twenty-one hours afterwards. The bargee had been away all the time, and we hold that his fabricated story was affirmative evidence that he had no excuse for his absence. At the locus in quo—especially during the short January days and in the full tide of war activity—barges were being constantly "drilled" in and out. Certainly it was not beyond reasonable expectation, that with the inevitable haste and bustle, the work might not be done with adequate care. In such circumstances we hold—and it is all that we do hold—that it was a fair requirement that the Conners Company should have a bargee aboard (unless he had some excuse for his absence), during the working hours of daylight.

REVIEW AND DISCUSSION QUESTIONS

1. Describe the facts in this case, along with the legal issue that Judge Hand is weighing.
2. Explain the importance of the formula $B < PI$ Does it suffice to determine whether it is fair to hold a particular defendant liable?
3. This case is often described as one that illustrates the "economic" approach to law. Why might somebody say that?
4. Hand does not hold that a bargee must always be on board. Why not?

Foreseeable Risk

Stone v. Bolton

Negligence, as William Prosser explained earlier, is conduct that falls below a standard established by the law for the protection of others against unreasonable risk of harm. The idea of risk, Prosser says, necessarily involves a recognizable danger: "A risk is a danger which is apparent, or should be apparent, to one in the position of the actor." But "the culpability of the actor's conduct must be judged in the light of the possibilities apparent to him at the time, and not by looking backward 'with wisdom born of the event.'" At issue in the English case of *Stone v. Bolton* is what exactly constitutes a reasonably foreseeable risk.

Stone v. Bolton, 1 K.B. 201 (C.A. 1950); *Bolton v. Stone*, A. C. 850 (1951).

[The plaintiff, Miss Bessie Stone, lived on Beckenham Road, a side street next to a cricket ground. One day, as she had just passed outside the gate in front of her house, she was struck on the head by a cricket ball that had been hit from the grounds. The ball was hit by a player on the visiting team, and by all accounts, was one of the longest balls—travelling about 100 yards before it struck the plaintiff—that had ever been hit at the grounds during the last 40 years. The cricket ground was found at trial to be "quite large enough for all practical purposes," even after it was remodeled in 1910 or 1911 to allow for construction of Beckenham Road. The field itself was surrounded by a twelve-foot-high fence or hoarding which, owing to a rise in the ground, was about seventeen feet above the street on the Beckenham Road side. The southern wicket from which the ball was struck was about 78 yards from Beckenham Road fence. Witnesses testified that over a 30-year period about six to ten balls had been hit onto Beckenham Road, and that several others had landed in the garden of one Mr. Brownson, whose house was the closest to the cricket grounds of all the houses in the neighborhood. The plaintiff did not sue the batsman or his club; but she did sue the home cricket club and all of its members, alleging first that the grounds constituted a public nuisance. A second and separate count was based on common law negligence. The particulars of negligence were that the defendants had placed the cricket pitch too close to Beckenham Road, that they failed to erect a fence of sufficient height to prevent balls from being hit onto the road, and that they otherwise failed to insure that cricket balls would not be hit into said road. At trial, Oliver, J., gave judgment to the defendants on both the public nuisance and negligence counts. That judgment was reversed by a two-to-one vote in the Court of Appeal, which held for the plaintiff on the negligence question.]

Jenkins, L. J. [for the Court of Appeal]: The case as regards negligence, therefore, seems to me to resolve itself into the question whether, with the wickets sited as they were, and the fence at the Beckenham Road end as it was, on August 9, 1947, the hitting into Beckenham Road of the ball which struck and injured the plaintiff was the realization of a reasonably foreseeable risk, or was in the nature of an unprecedented occurrence which the defendants could not reasonably have foreseen.

On the evidence this question seems to me to admit of only one answer. Balls had been hit into Beckenham Road before. It is true this had happened only at rare intervals, perhaps no more than six times in thirty seasons. But it was known from practical experience to be an actual possibility in the conditions in which matches were customarily played on the ground from about 1910 onwards, that is to say, with the wickets sited substantially as they were, and the fence at the Beckenham Road end, I gather, exactly as it was as regards height and position on August 9, 1947. What had happened several times before could, as it seems to me, reasonably be expected to happen again sooner or later. It was not likely to happen often, but it was certainly likely to happen again in time. When or how often it would happen again no one could tell, as this would depend on the strength of the batsmen playing on the ground (including visitors about whose capacity the defendants might know nothing) and the efficiency or otherwise of the bowlers. In my opinion, therefore, the hitting out of the ground of the ball which struck and injured the plaintiff was a realization of a reasonably foreseeable risk, which because it could reasonably be foreseen, the defendants were under a duty to prevent.

The defendants had, in fact, done nothing since the rearrangement of the ground on the making of Beckenham Road in or about 1910, whether by heightening the fence (e.g., by means of a screen of wire netting on poles) or by altering the position of the pitch, to guard against the known possibility of balls being hit into Beckenham Road. It follows that, if I have rightly defined the extent of the defendants' duty in this matter, the hitting out of the ground of the ball which injured the plaintiff did involve a breach of that duty for the consequences of which the defendants must be held liable to the plaintiff in damages. . . .

It was also, I think, suggested that no possible precaution would have arrested the flight of this particular ball, so high did it pass over the fence. This seems to me an irrelevant consideration. If cricket cannot be played on a given ground without foreseeable risk of injury to persons outside it, then

it is always possible in the last resort to stop using that ground for cricket. The plaintiff in this case might, I apprehend, quite possibly have been killed. I ask myself whether in that event the defendants would have claimed the right to go on as before, because such a thing was unlikely to happen again for several years, though it might happen again on any day on which one of the teams in the match included a strong hitter. No doubt as a practical matter the defendants might decide that the double chance of a ball being hit into the road and finding a human target there was so remote that, rather than go to expense in the way of a wire screen or the like, or worse still abandon the ground, they would run the risk of such an occurrence and meet any ensuing claim for damages if and when it arose. But I fail to see on what principle they can be entitled to require people in Beckenham Road to accept the risk, and, if hit by a ball, put up with the possibly very serious harm done to them as damnum sine injuria, unless able to identify, trace, and successfully sue the particular batsman who made the hit.

[After judgment went against the defendants in the Court of Appeal, they appealed their case to the House of Lords, which unanimously decided the case in their favor.]

Lord Reid: This case therefore raises sharply the question what is the nature and extent of the duty of a person who promotes on his land operations which may cause damage to persons on an adjoining highway. Is it that he must not carry out or permit an operation which he knows or ought to know clearly can cause such damage, however improbable that result may be, or is it that he is only bound to take into account the possibility of such damage if such damage is a likely or probable consequence of what he does or permits, or if the risk of damage is such that a reasonable man, careful of the safety of his neighbor, would regard that risk as material? . . .

Counsel for the respondent [Stone] in this case had to put his case so high as to say that, at least as soon as one ball had been driven into the road in the ordinary course of a match, the appellants [cricket ground] could and should have realized that that might happen again and that, if it did, someone might be injured; and that that was enough to put on the appellants a duty to take steps to prevent such an occurrence. If the true test is foreseeability alone I think that must be so. Once a ball has been driven on to a road without there being anything extraordinary to account for the fact, there is clearly a risk that another will follow, and if it does there is clearly a chance, small though it may be, that someone may be injured. On the theory that it is foreseeability alone that matters it would be irrelevant to consider how often a ball might be expected to land in the road and it would not matter whether the road was the busiest street, or the quietest country lane; the only difference between these cases is in the degree of risk.

It would take a good deal to make me believe that the law has departed so far from the standards which guide ordinary careful people in ordinary life. In the crowded conditions of modern life even the most careful person cannot avoid creating some risks and accepting others. What a man must not do, and what I think a careful man tries not to do, is to create a risk which is substantial. Of course there are numerous cases where special circumstances require that a higher standard shall be observed and where that is recognized by the law. But I do not think that this case comes within any such special category. . . . In my judgment the test to be applied here is whether the risk of damage to a person on the road was so small that a reasonable man in the position of the appellants, considering the matter from the point of view of safety, would have thought it right to refrain from taking steps to prevent the danger.

In considering that matter I think that it would be right to take into account not only how remote is the chance that a person might be struck but also how serious the consequences are likely to be if a person is struck; but I do not think that it would be right to take into account the difficulty of remedial measures. If cricket cannot be played on a ground without creating a substantial risk, then it should not be played there *at all.* . . . This is a question not of law but of fact and degree. It is not an easy question and it is one on which opinions may well differ. I can only say that having given the whole matter repeated and anxious consideration I find myself unable to decide this question in favour of the respondent [Stone]. But I think that this case is not far from the borderline. If this appeal is allowed, that does not in

my judgment mean that in every case where cricket has been played on a ground for a number of years without accident or complaint those who organize matches there are safe to go on in reliance on past immunity. I would have reached a different conclusion if I had thought that the risk here had been other than extremely small, because I do not think that a reasonable man considering the matter from the point of view of safety would or should disregard any risk unless it is extremely small.

REVIEW AND DISCUSSION QUESTIONS

1. Why does appellate judge Jenkins hold that Miss Stone's injury was a "reasonably foreseeable risk," rather than "an unprecedented occurrence which the defendants could not reasonably have foreseen"? Do you agree? What facts are the most relevant and important for answering this question?
2. How does Justice Jenkins respond to the claim that no possible precaution could have stopped this particular ball? If this claim is true, should the cricket grounds be closed?
3. In reversing the appellate court's decision, what factors does Lord Reid consider besides the foreseeability of risk?
4. Lord Reid argues that the appellate court's decision departs "from the standards which guide ordinary people in ordinary life." Do you agree? What would the reasonable man test tell us in this case?
5. If Lord Reid is correct, then Miss Stone apparently has no legal redress for the injury she has suffered. Is this fair? What would the Hand formula say?

Tort Liability and Corrective Justice

Jeffrie G. Murphy and Jules L. Coleman

When a person is injured, the initial legal presumption is that the injured person should bear the losses himself. As Jeffrie G. Murphy and Jules L. Coleman explain, the rules of tort liability spell out when a person's losses can be shifted to another party. Sometimes the rule of fault liability applies, at other times, the rule of strict liability. Reasons can be advanced for either fault or strict liability, and reflecting on these leads to questions about the whole nature of tort liability. Murphy and Coleman pursue these questions in the context of the debate over no-fault automobile insurance plans. A pure no-fault system would require drivers to carry insurance for their own protection but bar them from suing another driver for damages. By thus abandoning the tort system in accident cases, it would, its proponents believe, make auto insurance much cheaper. Whereas no-fault accident systems are defended on grounds of social utility, the fault system is generally upheld by appeals to retributive justice and compensatory or corrective justice. The authors argue, however, that when a tort is not intentional, neither retributive justice nor corrective justice require compensation to be based on fault. Jeffrie Murphy teaches philosophy at Arizona State University, and Jules Coleman teaches law and philosophy at Yale University.

THE RULES OF TORT LIABILITY

Often when an individual is injured by the doings of another, he suffers measurable damages. The law of torts provides the victim with a possible source of recovery. Were there no remedy, the victim would have to shoulder the costs occasioned by the conduct of others. In torts, we capture this by the expression that *initially* losses shall lie where they have fallen. In short, we begin with the presumption that victims should bear their losses. This is a presumption that can be overridden. Specifying the conditions under which the presumption can be overridden make up the body of tort law.

One reason we might have for our initial presumption that the victims should bear their losses is that it is costly to shift them to others. If there is no good reason for shifting a loss, why incur unnecessary costs? By letting the loss lie where it falls, in effect we hold the victim liable for the costs of harms he suffers, unless he or she is prepared to take steps to have us consider shifting the loss to another—usually but not always, his injurer. The victim has the burden of coming forward and initiating litigation and the additional burden of persuading a jury that his loss ought to be borne by another.

The courts recognize only some reasons as good ones for shifting losses or, in other words, for overriding the presumption that the victim should bear his own costs. These reasons are captured in *liability rules*. Liability rules set forth the conditions a victim must meet in order to have his losses borne by another. Usually these rules are not conclusive. If the conditions set out by a liability rule are satisfied, the burden is shifted from the victim to the defendant. The plaintiff-victim has met his burden; he has given the court good reason for thinking that he ought not bear his losses, and for thinking that the defendant should. Now the defendant has available to her the opportunity to present an argument to the effect that she ought not shoulder the victim's costs. Not every reason she might offer is a good one. The good ones are captured under the rules of *positive defense*.

We can illustrate the mechanisms involved [with an example]. . . . The home owner whose kitchen your car has entered will bear the costs your recklessness caused unless he can convince a jury that you should. In an effort to convince someone that the responsibility for these costs lies with you, he initiates litigation. He goes to court and establishes that you were reckless in your driving and that your recklessness caused his damages. He has met the initial burden by showing that you were at fault and that your fault caused the harm. His costs will be your responsibility unless you have a compelling reason why they should not be. Suppose it turns out that you drove haphazardly because earlier that night at a party the victim without your knowledge had slipped you severely intoxicating beverages. Moreover, the effects of the beverages would be delayed so that you would experience intoxication long after you started driving. The net effect is that you would be unable to reason clearly about what you ought to do in your intoxicated state. All would be beyond your control. In that case, you would have shown that the victim's very own mischief was responsible for the damage he suffered. In doing so, you have offered a positive defense sufficient to shift the burden back to him. Not every defense you might have tried would have been so successful. Had you argued that you were drunk because you are an alcoholic, you might have received some sympathy from the court, but you could hardly have expected to be freed of the burden of recompense.

In this section, we will focus on liability rules rather than on the rules of positive defense. In other words, we will be concerned with the conditions that must be satisfied in order for a victim to overcome the initial presumption that the losses that have befallen him should remain his responsibility. In general, a distinction is drawn between two kinds of liability rules; these are the rules of fault and of strict liability. If a case is covered by a rule of strict liability, the victim must establish that:

1. The defendant acted.
2. He, the victim, suffered a compensable loss.
3. The defendant's action caused the loss.

Under a rule of fault liability, the plaintiff-victim must establish that each of these three conditions has been satisfied and that the following condition has been satisfied as well.

4. The defendant's conduct was at fault.

Under the rule of strict liability, it is enough that the plaintiff show that the injurer caused the harm for which he seeks recompense. Under the rule of fault liability, he must show both that the injurer is at fault and that his fault is causally responsible for the victim's loss.[1] Certain cases that come before a court are decided on a strict liability basis, others on a fault basis. One question in tort theory is which sorts of cases should be adjudicated under which sort of rule. This question presupposes that we have already determined the relative advantages of each sort of liability rule, and what we are trying to do is to fit the rules to their most

appropriate applications. The philosopher of law wants to look more closely at what the traditional analysis takes for granted; he or she wants to know what reasons can be advanced on behalf of either strict or fault liability. The legal philosopher wants to know, in other words, what the relative virtues of the rules are. Does fault liability, for example, promote justice, and strict liability, utility? Or is it the other way around? Indeed, the legal philosopher is troubled by an even more basic question: why rectify gains and losses according to the traditions, customs, and rules developed over the history of tort law? Maybe the whole system of liability rules and positive defence fails to achieve either justice or utility—or maybe it is less successful at promoting justice or utility than alternatives to it might be.

In the law of torts, individuals bring their claims to court on a case-by-case basis. But, theoretically, we could deal with their claims in other ways, for example, by lumping together the costs of certain kinds of harms as they accumulate among victims over a period of time, and then compensate the victims as a class through the tax coffers. I don't want to claim that such a plan is preferable to the tort system, but I do want to suggest that we cannot determine if it is unless we first determine what goals we think the law of torts ought to pursue. Only when we have an idea of what it is we want the law of torts to accomplish, will we be able to determine if all cases should be adjudicated on a case-by-case basis. Only then will we be able to determine the relative advantages of the rules of strict and fault liability in securing them.

I want now to explore these questions in the context of a particular issue within the law of torts: namely, the debate over no-fault automobile insurance. No-fault presents a good focal point for this discussion because under a no-fault system, cases are not, in fact, adjudicated individually; indeed, in theory they are not to be adjudicated at all. No-fault, in short, involves abandoning the tort law system as we know it. The costs of accidents under a no-fault scheme are not distributed according to the rule of fault liability, and this seems initially troubling to those who think that justice requires imposing the costs of accidents upon those whose fault occasions them. The intuitive view is that the fault liability rule promotes justice, and that the rule of strict liability abandons justice to promote utility. Abandoning the tort system in general and the rule of fault liability in particular as instruments for rectifying wrongful losses that result from negligent motoring seems to conflict with our ideals of just law, and in what follows I want to see if this initial impression stands up under scrutiny.

JUSTICE AND THE COSTS OF AUTOMOBILE ACCIDENTS

. . . Interestingly enough, if justice were to require that fault-related burdens be distributed exclusively according to the criterion of fault, it would be very hard to imagine how such demands could be met by the fault system, at least as we know it. In the first place, all traditional forms of insurance would be ruled out since they would allocate burdens in part among individuals whose records have been (and may remain) spotless, as well as among past and present wrongdoers. The burdens of the faultless would not reflect their desert and would be objectionable on those grounds. In the absence of insurance, the costs of particular accidents would fall either on individual victims or injurers, with victims being ruled out on the same grounds as the faultless: no fault, no penalty. But if the costs of accidents were to fall entirely on individual wrongdoers, they would often prove disproportionate to the wrongdoing involved—slight errors can result in costly accidents—and this would be objectionable on the grounds that the burden would not reflect the degree of fault. . . .

. . . [D]efenders of the fault system [believe] that distributing losses in this way—by having them fall initially on individual injurers—secures the goal of justice, a legitimate aim of accident law, by penalizing the wrongdoing of those at fault. The fault system, by penalizing the wrongful conduct, is therefore required by principles of retributive justice. On related grounds, prohibitions against no-fault are supported by principles of compensatory justice, which are said to require that recompense for losses must originate with those at fault. I want to consider both of these arguments in turn.

According to the standard conception of retributive justice, wrongdoing deserves its comeuppance: a measure of pain, suffering, or deprivation

should be exacted from wrongdoers, and the deprivation should reflect the nature and magnitude of the wrongdoing. In its most prevalent and, I believe, least acceptable form the retributivist principle is a moral position in a very strong sense that in order for penalization to be deserved, the defective conduct must be morally defective, and the author of it morally at fault. Thus legal fault, insofar as it is based on the moral fault of the actor, ought to be penalized.

One of the most prevalent defenses of a fault system rests on the claim that this principle of retributive justice requires it. The argument usually takes the following form. If an actor's defective conduct is a substantial contributory factor in bringing about harm, we impute the harm to him as his fault. He has, in the legal jargon, satisfied the fault requirement. The fault system, by making those at fault in causing accidents liable for their costs, guarantees that culpable motoring gets its comeuppance and is therefore required by the retributivist principle.

This argument in favor of the fault system does not purport to establish that every time an actor is at fault he is morally blameworthy. It establishes only that he is blameworthy in some sense of the term. And although some faulty driving is morally defective, for example, drunken, reckless driving, most driving faults are not moral shortcomings. If the retributivist principle were to support penalties only for moral wrongdoing, the fault system would justify too much. However, we need not interpret the retributivist principle this narrowly. Perhaps we should understand it as endorsing penalties for nonmoral as well as for moral shortcomings. H. L. A. Hart, for one, has argued persuasively that it is permissible morally to penalize wrongdoing that is an actor's fault—to his genuine discredit—whether or not the flaw in his conduct is a moral one.[2] If Hart is correct, then there exists a moral license to impose penalties on drivers who fail to satisfy the legal standard of due care, provided that failure can be charged against their personal records, that is, is something for which they are to blame. Hart's position does not amount to an endorsement of retributivism since he claims only that genuine wrongdoing may be penalized, while retributivists claim that such penalties are not merely permissible but a matter of right action.

Being at fault, morally or otherwise, has two essential ingredients. One element in fault judgments relates to the character of the act and its relationship to the appropriate standard of conduct, while the other relates to the actor's "state of mind." Most fault judgments are true just in case both the act-regarding and the actor-regarding requirements are satisfied. The essence of these requirements is perhaps best illustrated by the defenses alleged wrongdoers offer to defeat accusations of fault—that is, justifications and excuses. When an actor has a justification for what he has done, he denies that the act, all things considered, fails to satisfy the appropriate norm, or he argues that his conduct is an exception to the rule usually governing that type of conduct, or that it is an exception to the rule that such conduct is wrong. Should he have an excuse for what he has done, he would not be denying that his conduct was substandard, only that it was his fault. The "subjective" or mental element necessary to charge the act against his record, he thus argues, is missing.

If an actor has an excuse for his actions, he is not at fault in the sense appropriate to justify imposing penalties against him on the grounds of retributive justice. That is because retributivism licenses penalties for *genuine* wrongdoing only. Insofar as excuses evidence the nonvoluntary character of the conduct, failure to measure up under excusing conditions cannot be considered something for which the actor is to blame. Though the retributivist principle would not endorse penalties for what would ordinarily count as excusable departures from the standard of due care, the system of awarding liability for accident costs on the basis of fault often penalizes those who in Hart's words "could not help doing what they did."

This feature of the fault system is well known among lawyers and has led some legal theorists to question if the courts should apply subjective (internal) or objective (external) standards of fault to determine liability in accident law. In other words, should the law require no more of each member of the community than he is capable of (the subjective test of fault) or should it require each to live up to a standard that may exceed the capacities of some (the objective test)? The most

famous and perhaps still the most compelling defense of the objective test is articulated in Oliver Wendell Holmes:

> If, for instance, a man is born hasty and awkward, is always having accidents and hurting himself or his neighbors, no doubt his congenital defects will be allowed for in the courts of Heaven, but his slips are no less troublesome to his neighbors than if they sprang from guilty neglect. His neighbors, accordingly require him, at his proper peril, to come up to their standard, and the courts which they establish decline to take his personal equation into account.[3]

Ascriptions of fault that are verified by applying objective criteria of conduct are not defeasible by excuses. That is their distinctive feature. Thus, if an injurer genuinely could do no better than he did—if, in other words, his conduct was, in a suitably narrow sense, nonvoluntary—this would *not* suffice to free him from the burdens of recompense under the present system of accident law, though the nonvoluntary character of his actions would normally constitute excusing conditions and free him from blame.

We could of course amend the fault system to obviate this sort of difficulty by requiring that it apply only subjective criteria of fault. In this way, the fault system would penalize only those whose failure to exercise due care is to their personal discredit. But even if we recommend the application of the subjective test, the fault system would fail to satisfy the retributivist principle. That is because, in the fault system, only accident-causing wrongful driving receives its due—where the penalty is liability for accident costs—leaving unpenalized the class of dangerous and unnecessarily risky, but not harm-causing, wrongful motoring. From the retributivist point of view, this is unacceptable, since suitably culpable conduct without harmful causal upshots falls within its ambit. Thus, in the criminal law, so-called wrongful attempts are punishable, for example, attempted murder, attempted rape, and so on. From the retributivist point of view, the punishment of wrongful attempts is mandatory. The rough civil law analogue of criminal attempts is conduct that is "at fault," that is, unreasonably risky, harm-threatening, or dangerous. If securing the goal of retributive justice were the only goal of accident law, and if the penalty of liability for accident costs were the only sort of penalty befitting wrongful conduct of this sort, securing that goal would require that liability be spread among those at fault, not merely among those at fault in causing injuries. But even in this system driving faults are rendered market, not moral, values.

Moreover, the retributivist principle requires only that a fitting penalty be administered to wrongdoers. Nothing in the principle specifies that the penalty be liability for accident costs. Indeed, we have seen some of the difficulties inherent in trying to "fit" the penalty of civil liability to driving faults. An accident law in which liability for accident costs is the penalty for wrongful driving conflates two distinguishable issues: those of retribution and recompense. The fault system may be characterized by the fact that it joins these two issues, since in that system, the retribution for wrongdoing is recompense for injuries that are one's fault. The distinguishing feature of no-fault is the separation of these two concerns. By referring to accident law as a no-fault system, we do not mean to imply that there exist no purposes for which conduct should be examined for fault. Indeed, the criterion of fault is perfectly relevant to contriving penal fines or other penalties themselves suggested by the retributivist principle: penalties more fitting than civil liability in the sense that they are not subject to market fluctuations.[4] All we mean to imply by the label "no-fault" is that the criterion of fault is irrelevant to the issue of recompense. That goal is to be satisfied by ensuring that the traffic victim receives the compensation to which he is entitled. Our notions of retributive justice therefore do not require that compensation, in order to be just, originate with those at fault. Whether or not our notions of corrective justice do is what I now want to consider.

There are a number of statements or ways of characterizing the requirements of corrective justice. I want to begin with the following conception of corrective justice. On this view, corrective or rectificatory justice is concerned with wrongful gains and losses. Rectification is, on this view, a matter of justice when it is necessary to protect a distribution of holdings (or entitlements) from distortions that arise from unjust enrichments or wrongful losses. The principle of corrective justice

requires the annulment of both wrongful gains and losses.

In order to invoke the principle of corrective justice to support eliminating or rectifying a distortion in a distribution of holdings or entitlements, the distribution need not itself be just. Corrective justice is a matter of justice on this view *not* because it promotes justice in the distribution of holdings, but rather because it remedies unjust departures from the prevailing distribution of holdings. However, following the requirements of corrective justice is necessary to maintain a just distribution of holdings. Therefore, any theory of distributive justice must make provisions for a theory of corrective justice. Corrective justice is an independent principle of justice precisely because it may be legitimately invoked to protect or reinstate distributions of holdings that would themselves fail the test of distributive justice.

Under the principle of corrective or compensatory justice, the victim of another's fault may have a right to recompense. The principle of corrective justice does not establish the right of victims to recompense for "faultlessly" caused injuries, and though an insurance scheme that protected policy holders against errorless as well as faultily caused injuries might be preferable, the protection such a scheme could offer would extend beyond what injured parties could claim as their *right*.

If compensation for wrongfully inflicted harms is one's right, against whom does the right bearer hold it? Upon whom does the corresponding obligation fall? In an insurance scheme, the injured party's right to recompense may be against his insurance company or against the insurance company of his injurer. That depends on whether the insurance coverage is a "first" or "third" party plan; that is, whether individuals insure themselves against personal loss owing to the conduct of others (or themselves)—first party—or against liability for the injuries they may cause others—third party. According to the most extensive form of no-fault coverage, the victim's right to recompense is a valid claim he holds against his own insurance company: that right being derived from contract. The underlying moral issue is whether or not such contracts and their corresponding rights and obligations are consistent with principles of corrective justice. Is compensation for accidental harms to be analyzed broadly to require only that victims receive the compensation to which they are entitled, or narrowly to require that recompense for victims must originate with respective injurers?

Suppose now that two individuals, X and Y, were involved in an accident that is Y's fault. X's right to compensation creates a corresponding duty to provide recompense that seems to fall on Y, his injurer. It seems, moreover, that justice would require that Y compensate X. If this is so, it suggests the general conclusion that, in order to be just, compensation must come from those at fault.

Why is it that justice requires that Y bear X's costs? Perhaps, if Y does not compensate X, X will go without compensation. One consequence of this would be that X's conduct would be penalized. Certainly that would be unjust. But that justice is avoidable by guaranteeing that X does not have to bear his own costs. That goal is accomplished when anyone other than X, including, but not necessarily, Y shoulders them. Perhaps by Y not bearing X's costs, Y's wrongdoing would go unpenalized. That would presumably constitute a retributive injustice. But the retributive principle requires only that Y's wrongdoing be penalized, and that goal is accomplished not only when Y is made to compensate X, but also when Y is penalized in some other way, for example, through penal channels.

Both compensation and punishment are concerned with wrongdoing—wrongful gain, advantage, or benefit. But where punishment involves wrongdoers incurring some evil—usually the loss of assorted legal rights—for their wrongfully obtained (or sought) gains or advantages, compensation aims at annulling, rectifying, or eliminating these undeserved or otherwise unjustifiable gains. Where punishment is concerned with victims only secondarily, the overriding concern of compensation is the nullification of the victim's losses; the reordering of his affairs to make him whole again. James Nickel has simply but insightfully captured this feature of compensatory justice as the elimination of unjustifiable *gains and losses* owing to human action—what he aptly terms "distortions."[5]

If compensation involves the elimination of distortions, surely we must acknowledge that not every means of securing that goal would be

endorsed by principles of compensatory justice. However seductive the myth of Robin Hood may be, there is little evidence to suggest that such a method for nullifying unwarranted gains and losses would satisfy the demands of compensatory justice. What further distinguishes compensatory from other principles of justice is that, in order to secure the elimination of distortions, it supports a system of correlative rights and duties between respective victims and wrongdoers. In the typical case of compensation, finding for the plaintiff amounts to a recognition of his or her (legal) rights to recompense; a right that imposes a correlative duty upon the defendant to provide compensation. In the typical instance of punishment, a verdict of guilty confers upon the state a right to impose some penalty against the defendant. But this right is a moral license, and unlike a claim-right, it does not give rise to a correlative duty; in this case, an obligation on the defendant to be punished or to allow himself to be punished.

But if a distinguishing feature of compensatory justice is this structure of correlative rights and duties, wouldn't principles of compensatory justice therefore prohibit contractual relations that supersede this structure by imposing duties of recompense on the *victim's* insurance company—contractual relations that require the victim's insurance company to discharge what are, in fact, the duties of wrongdoers? Compensation is a kind of repayment, and it is therefore with the category of obligations to repay that we are concerned. In the case of obligations to repay debts or loans, justice does not as a rule require that the repayment originate with the party obligated to repay. Thus, when someone borrows money from the local bank, few protests of injustice are heard above clicking safe locks, provided the money is returned in sufficient quantities and at appropriate intervals, regardless of exactly who is forwarding the payments on whose behalf. My obligation to repay a debt of, say, $100 to you is discharged when my wealthy patron approaches you with that sum, and you accept it as repayment, even if I know nothing of this transaction. Of course some obligations to repay may be described so determinately that they could not justly be discharged by another. Here the requirement that the repayment originate with the obligated party is built into the description of the obligation, and the fact that it cannot justly be discharged by another is a function of the determinacy of the obligation, not of justice.

Though it is a misleading way of talking, we sometimes describe punishment as a kind of repayment. We talk, for example, about criminals as owing "debts to society," which are presumably repaid through incarceration. If there is any merit in viewing punishment as a debt of repayment, then surely it is the sort of debt that in order to be just must be repaid by the "obligated" party, that is, the wrongdoer.[6] Now, is the obligation to provide compensation for accidental torts something that can justly be discharged by intervening parties—in this case, insurance companies—and therefore like obligations to repay debts in this respect, or is it something that can be discharged justly only by wrongdoers—like the purported obligation criminals have to pay back society? If compensation is, in this regard, like punishment, in order to be just, recompense for accidental torts must therefore originate with those at fault.

Unlike the obligation to repay debts, the obligation to compensate traffic victims is not derived from contractual relation or promise. It is derived from wrongdoing. But wrongdoing plays significantly different roles in punishment and in compensation. In compensation, unlike punishment, wrongdoing is not viewed as something in itself worthy of penalty. Instead compensation is concerned with wrongdoing only insofar as it either involves wrongful gain or benefit at another's expense, or it is evidence of the unjustifiable character of the victim's loss. In the absence of wrongful gain, proof of wrongdoing supports the victim's assertion that his losses have been wrongfully absorbed, and therefore justifies his demanding compensation as a matter of right, rather than requesting it as a matter of benevolence, utility, or welfare.

The distinguishing feature of automobile accidents and of accidental torts generally is that the injurer does not, as a rule, gain from his wrong, nor is it plausible to interpret his conduct as directed at securing some temporary or long-term

advantage. It is, after all, the nature of the beast; accidents are unintentional. One does not plan an accident. On the contrary, intentional torts such as fraud are undertaken to secure wrongful advantage at another's expense. A no-fault allocation of losses owing to intentional torts involving gain would not nullify the wrongdoer's gain and would be objectionable on those grounds. Nor would punishing him under the criminal law eliminate the gain, since the penal statute requires only that he suffer some evil for his wrong; nothing in it requires that he forfeit his gain. But in the case of accidental torts there is, in general, no gain on the wrongdoer's behalf that needs to be eliminated. That his conduct is wrongful supports the right of the victim to recompense, nothing more. Of course, his conduct may exhibit sufficient fault to merit penalty, but that is a concern of retributive justice and may best be dealt with by penal fine. Thus, it is my contention that principles of compensatory justice would not support a no-fault allocation of losses owing to certain intentional torts—i.e., because of the element of wrongful gain—but would not prohibit no-fault accident law precisely because of the absence of wrongful gain.[7]

This conclusion should be especially comforting to legal reformers such as Guido Calabresi and Jeffrey O'Connell, both of whom have recently urged the separation of accident law from the main body of tort law. Such proposals suggest that accident law govern recompense for harms owing primarily to negligence and defective products, and that they do so on a no-fault, no-defect basis. They are usually recommended to us by dint of their cost-saving features. The arguments I have been considering in this section on compensatory justice suggest still further support for their position. It is hoped that support from these quarters will surprise only those who are wedded to the view that justice and utility cannot be achieved except at one another's expense.

NOTES

1. A person is at fault in torts if his conduct is either negligent, reckless, or is the result of a wrongful intent to injure. A negligent person is one who simply fails to exercise the care of a reasonable man or woman of ordinary prudence.
2. Cf. H. L. A. Hart, "Legal Responsibility and Excuses," in *Punishment and Responsibility* (New York: Oxford, 1968), 39.
3. O. W. Holmes, *The Common Law* (Boston: Little, Brown and Company, 1963), 86.
4. Though such penalties are not affected by market values, they may be subject to considerations other than moral fault, e.g., deterrent value.
5. James Nickel, "What Is Compensatory Justice?" *William and Mary Law Review*, Vol. 3.
6. Imagine the effects on our notions of responsibility of an insurance scheme for criminal liability!
7. What I want to claim here is not that we can't imagine any situations in which accidents are beneficial to wrongdoers, but only that accidents in general are not beneficial to them, and, moreover, that accidents in general do not discriminate between victims and injurers in terms of likelihood of harm or gain. In other words, a priori, there is no reason to think that injurers more than victims will gain from accidents, just as there is no reason to assume that victims more than injurers will lose from them.

REVIEW AND DISCUSSION QUESTIONS

1. What is the difference between fault liability and strict liability? What is no-fault automobile insurance, and why would such a system involve "abandoning the tort law system as we know it"?
2. Fault has two essential ingredients. What are they, and how do they relate to justification and excuse? How does the contrast between subjective (internal) and objective (external) standards of fault apply to automobile accident cases?
3. Assess the argument that retributive justice requires the fault system in automobile accident cases.
4. What is corrective justice? Are Murphy and Coleman right that it does not require the fault system? Explain.

Negligence and Due Care

Palsgraf v. Long Island Railroad Co.

In this very famous and closely reasoned case, the New York Court of Appeals debates the nature and limits of negligence as a tort. Justice Benjamin Cardozo, one of the most distinguished jurists of the twentieth century and later a Supreme Court justice, contends that the negligence of the defendant's guards was not a wrong in relation to Mrs. Palsgraf, even though it resulted in her being injured. In his dissent, Judge Andrews rejects the idea that negligence must involve the breach of a duty owed to a particular person. He argues, rather, that one is liable for all the "proximate" consequences of actions that threaten the safety of others.

Chief Justice Cardozo: Plaintiff was standing on a platform of defendant's railroad after buying a ticket to go to Rockaway Beach. A train stopped at the station, bound for another place. Two men ran forward to catch it. One of the men reached the platform of the car without mishap, though the train was already moving. The other man, carrying a package, jumped aboard the car, but seemed unsteady as if about to fall. A guard on the car, who had held the door open, reached forward to help him in, and another guard on the platform pushed him from behind. In this act, the package was dislodged, and fell upon the rails. It was a package of small size, about fifteen inches long, and was covered by a newspaper. In fact it contained fireworks, but there was nothing in its appearance to give notice of its contents. The fireworks when they fell exploded. The shock of the explosion threw down some scales at the other end of the platform, many feet away. The scales struck the plaintiff, causing injuries for which she sues.

The conduct of the defendant's guard, if a wrong in its relation to the holder of the package, was not a wrong in its relation to the plaintiff, standing far away. Relatively to her it was not negligence at all. Nothing in the situation gave notice that the falling package had in it the potency of peril to persons thus removed. Negligence is not actionable unless it involves the invasion of a legally protected interest, the violation of a right. "Proof of negligence in the air, so to speak, will not do." "Negligence is the absence of care, according to the circumstances." The plaintiff as she stood upon the platform of the station might claim to be protected against intentional invasion of her bodily security. Such invasion is not charged. She might claim to be protected against unintentional invasion by conduct involving in the thought of reasonable men an unreasonable hazard that such invasion would ensue. These, from the point of view of the law, were the bounds of her immunity, with perhaps some rare exceptions, survivals for the most part of ancient forms of liability, where conduct is held to be at the peril of the actor. . . . If no hazard was apparent to the eye of ordinary vigilance, an act innocent and harmless, at least to outward seeming, with reference to her, did not take to itself the quality of a tort because it happened to be a wrong, though apparently not one involving the risk of bodily insecurity, with reference to some one else. "In every instance, before negligence can be predicated of a given act, back of the act must be sought and found a duty to the individual complaining, the observance of which would have averted or avoided the injury." "The ideas of negligence and duty are strictly correlative." (Bowen, L. J., in *Thomas v. Quartermaine*, 18 Q. B. D. 685, 694). The plaintiff sues in her own right for a wrong personal to her, and not as the vicarious beneficiary of a breach of duty to another.

A different conclusion will involve us, and swiftly too, in a maze of contradictions. A guard stumbles over a package which has been left upon a platform. It seems to be a bundle of newspapers. It turns out to be a can of dynamite. To the eye of ordinary vigilance, the bundle is abandoned waste,

248 N.Y. 339 (1928). Some citations omitted.

which may be kicked or trod on with impunity. Is a passenger at the other end of the platform protected by the law against the unsuspected hazard concealed beneath the waste? If not, is the result to be any different, so far as the distant passenger is concerned, when the guard stumbles over a valise which a truckman or a porter has left upon the walk? The passenger far away, if the victim of a wrong at all, has a cause of action, not derivative, but original and primary. His claim to be protected against invasion of his bodily security is neither greater nor less because the act resulting in the invasion is a wrong to another far removed. In this case, the rights that are said to have been violated, the interests said to have been invaded, are not even of the same order. The man was not injured in his person nor even put in danger. The purpose of the act, as well as its effect, was to make his person safe. If there was a wrong to him at all, which may very well be doubted, it was a wrong to a property interest only, the safety of his package. Out of this wrong to property, which threatened injury to nothing else, there has passed, we are told, to the plaintiff by derivation or succession a right of action for the invasion of an interest of another order, the right to bodily security. The diversity of interests emphasizes the futility of the effort to build the plaintiff's right upon the basis of a wrong to some one else. The gain is one of emphasis, for a like result would follow if the interests were the same. Even then, the orbit of the danger as disclosed to the eye of reasonable vigilance would be the orbit of the duty. One who jostles one's neighbor in a crowd does not invade the rights of others standing at the outer fringe when the unintended contact casts a bomb upon the ground. The wrongdoer, as to them, is the man who carries the bomb, not the one who explodes it without suspicion of the danger. Life will have to be made over, and human nature transformed, before prevision so extravagant can be accepted as the norm of conduct, the customary standard to which behavior must conform.

The argument for the plaintiff is built upon the shifting meanings of such words as "wrong" and "wrongful," and shares their instability. What the plaintiff must show is "a wrong" to herself, *i.e.*, a violation of her own right, and not merely a wrong to some one else, nor conduct "wrongful" because unsocial, but not "a wrong" to any one. We are told that one who drives at reckless speed through a crowded city street is guilty of a negligent act and, therefore, of a wrongful one irrespective of the consequences. Negligent the act is, and wrongful in the sense that it is unsocial, but wrongful and unsocial in relation to other travelers, only because the eye of vigilance perceives the risk of damage. If the same act were to be committed on a speedway or a race course, it would lose its wrongful quality. The risk reasonably to be perceived defines the duty to be obeyed, and risk imports relation; it is risk to another or to others within the range of apprehension. . . . This does not mean, of course, that one who launches a destructive force is always relieved of liability if the force, though known to be destructive, pursues an unexpected path. It was not necessary that the defendant should have had notice of the particular method in which an accident would occur, if the possibility of an accident was clear to the ordinarily prudent eye. . . . Some acts, such as shooting, are so imminently dangerous to any one who may come within reach of the missile, however unexpectedly, as to impose a duty of prevision not far from that of an insurer. Even today, and much oftener in earlier stages of the law, one acts sometimes at one's peril. Under this head, it may be, fall certain cases of what is known as transferred intent, an act willfully dangerous to A resulting by misadventure in injury to B. These cases aside, wrong is defined in terms of the natural or probable, at least when unintentional. The range of reasonable apprehension is at times a question for the court, and at times, if varying inferences are possible, a question for the jury. Here, by concession, there was nothing in the situation to suggest to the most cautious mind that the parcel wrapped in newspaper would spread wreckage through the station. If the guard had thrown it down knowingly and willfully, he would not have threatened the plaintiff's safety, so far as appearances could warn him. His conduct would not have involved, even then, an unreasonable probability of invasion of her bodily security. Liability can be no greater where the act is inadvertent.

Negligence, like risk, is thus a term of relation. Negligence in the abstract, apart from things related, is surely not a tort, if indeed it is understandable at all. Negligence is not a tort unless it results

in the commission of a wrong, and the commission of a wrong imports the violation of a right, in this case, we are told, the right to be protected against interference with one's bodily security. But bodily security is protected, not against all forms of interference or aggression, but only against some. One who seeks redress at law does not make out a cause of action by showing without more that there has been damage to his person. If the harm was not willful, he must show that the act as to him had possibilities of danger so many and apparent as to entitle him to be protected against the doing of it though the harm was unintended. Affront to personality is still the keynote of the wrong. Confirmation of this view will be found in the history and development of the action on the case. Negligence as a basis of civil liability was unknown to mediaeval law. For damage to the person, the sole remedy was trespass, and trespass did not lie in the absence of aggression, and that direct and personal. Liability for other damage, as where a servant without orders from the master does or omits something to the damage of another, is a plant of later growth. When it emerged out of the legal soil, it was thought of as a variant of trespass, an offshoot of the parent stock. This appears in the form of action, which was known as trespass on the case. The victim does not sue derivatively, or by right of subrogation, to vindicate an interest invaded in the person of another. Thus to view his cause of action is to ignore the fundamental difference between tort and crime. He sues for breach of a duty owing to himself.

The law of causation, remote or proximate, is thus foreign to the case before us. The question of liability is always anterior to the question of the measure of the consequences that go with liability. If there is no tort to be redressed, there is no occasion to consider what damage might be recovered if there were a finding of a tort. We may assume, without deciding, that negligence, not at large or in the abstract, but in relation to the plaintiff, would entail liability for any and all consequences, however novel or extraordinary. There is room for argument that a distinction is to be drawn according to the diversity of interests invaded by the act, as where conduct negligent in that it threatens an insignificant invasion of an interest in property results in an unforeseeable invasion of an interest of another order, as *e.g.*, one of bodily security. Perhaps other distinctions may be necessary. We do not go into the question now. The consequences to be followed must first be rooted in a wrong.

The judgment of the Appellate Division and that of the Trial Term should be reversed, and the complaint dismissed, with costs in all courts.

ANDREWS, J. (dissenting): Assisting a passenger to board a train, the defendant's servant negligently knocked a package from his arms. It fell between the platform and the cars. Of its contents the servant knew and could know nothing. A violent explosion followed. The concussion broke some scales standing a considerable distance away. In falling they injured the plaintiff, an intending passenger.

Upon these facts may she recover the damages she has suffered in an action brought against the master? The result we shall reach depends upon our theory as to the nature of negligence. Is it a relative concept—the breach of some duty owing to a particular person or to particular persons? Or where there is an act which unreasonably threatens the safety of others, is the doer liable for all its proximate consequences, even where they result in injury to one who would generally be thought to be outside the radius of danger? This is not a mere dispute as to words. We might not believe that to the average mind the dropping of the bundle would seem to involve the probability of harm to the plaintiff standing many feet away whatever might be the case as to the owner or to one so near as to be likely struck by its fall. If, however, we adopt the second hypothesis we have to inquire only as to the relation between cause and effect. We deal in terms of proximate cause, not of negligence.

Negligence may be defined roughly as an act or omission which unreasonably does or may affect the rights of others, or which unreasonably fails to protect oneself from the dangers resulting from such acts. Here I confine myself to the first branch of the definition. Nor do I comment on the word "unreasonable." For present purposes it sufficiently describes that average of conduct that society requires of its members.

There must be both the act or the omission, and the right. It is the act itself, not the intent of the actor, that is important. In criminal law both

the intent and the result are to be considered. Intent again is material in tort actions, where punitive damages are sought, dependent on actual malice—not on merely reckless conduct. But here neither insanity nor infancy lessens responsibility.

As has been said, except in cases of contributory negligence, there must be rights which are or may be affected. Often though injury has occurred, no rights of him who suffers have been touched. A licensee or trespasser upon my land has no claim to affirmative care on my part that the land be made safe. Where a railroad is required to fence its tracks against cattle, no man's rights are injured should he wander upon the road because such fence is absent. An unborn child may not demand immunity from personal harm.

But we are told that "there is no negligence unless there is in the particular case a legal duty to take care, and this duty must be one which is owed to the plaintiff himself and not merely to others." This, I think too narrow a conception. Where there is the unreasonable act, and some right that may be affected there is negligence whether damage does or does not result. That is immaterial. Should we drive down Broadway at a reckless speed, we are negligent whether we strike an approaching car or miss it by an inch. The act itself is wrongful. It is a wrong not only to those who happen to be within the radius of danger but to all who might have been there—a wrong to the public at large. Such is the language of the street. Such is the language of the courts when speaking of contributory negligence. Such again and again their language in speaking of the duty of some defendant and discussing proximate cause in cases where such a discussion is wholly irrelevant on any other theory. As was said by Mr. Justice Holmes many years ago, "the measure of the defendant's duty in determining whether a wrong has been committed is one thing, the measure of liability when a wrong has been committed is another." Due care is a duty imposed on each one of us to protect society from unnecessary danger, not to protect A, B or C alone.

It may well be that there is no such thing as negligence in the abstract. "Proof of negligence in the air, so to speak, will not do." In an empty world negligence would not exist. It does involve a relationship between man and his fellows. But not merely a relationship between man and those whom he might reasonably expect his act would injure. Rather, a relationship between him and those whom he does in fact injure. If his act has a tendency to harm some one, it harms him a mile away as surely as it does those on the scene. We now permit children to recover for the negligent killing of the father. It was never prevented on the theory that no duty was owing to them. A husband may be compensated for the loss of his wife's services. To say that the wrongdoer was negligent as to the husband as well as to the wife is merely an attempt to fit facts to theory. An insurance company paying a fire loss recovers its payment of the negligent incendiary. We speak of subrogation—of suing in the right of the insured. Behind the cloud of words is the fact they hide, that the act, wrongful as to the insured, has also injured the company. Even if it be true that the fault of father, wife or insured will prevent recovery, it is because we consider the original negligence not the proximate cause of the injury.

In the well-known *Polemis* case, Scrutton, L. J., said that the dropping of a plank was negligent for it might injure "workman or cargo or ship." Because of either possibility the owner of the vessel was to be made good for his loss. The act being wrongful the doer was liable for its proximate results. Criticized and explained as this statement may have been, I think it states the law as it should be and as it is.

The proposition is this. Every one owes to the world at large the duty of refraining from those acts that may unreasonably threaten the safety of others. Such an act occurs. Not only is he wronged to whom harm might reasonably be expected to result, but he also who is in fact injured, even if he be outside what would generally be thought the danger zone. There needs be duty due the one complaining but this is not a duty to a particular individual because as to him harm might be expected. Harm to some one being the natural result of the act, not only that one alone, but all those in fact injured may complain. We have never, I think, held otherwise. Indeed in the *Di Caprio* case we said that a breach of a general ordinance defining the degree of care to be exercised in one's calling is evidence of negligence as to every one. We did not limit this statement to those who might be expected to be exposed to danger.

Unreasonable risk being taken, its consequences are not confined to those who might probably be hurt.

If this be so, we do not have a plaintiff suing by "derivation or succession." Her action is original and primary. Her claim is for a breach of duty to herself—not that she is subrogated to any right of action of the owner of the parcel or of a passenger standing at the scene of the explosion.

The right to recover damages rests on additional considerations. The plaintiff's rights must be injured, and this injury must be caused by the negligence. We build a dam, but are negligent as to its foundations. Breaking, it injures property down stream. We are not liable if all this happened because of some reason other than the insecure foundation. But when injuries do result from our unlawful act we are liable for the consequences. It does not matter that they are unusual, unexpected, unforeseen and unforeseeable. But there is one limitation. The damages must be so connected with the negligence that the latter may be said to be the proximate cause of the former.

These two words have never been given an inclusive definition. What is a cause in a legal sense, still more what is a proximate cause, depend in each case upon many considerations, as does the existence of negligence itself. Any philosophical doctrine of causation does not help us. A boy throws a stone into a pond. The ripples spread. The water level rises. The history of that pond is altered to all eternity. It will be altered by other causes also. Yet it will be forever the resultant of all causes combined. Each one will have an influence. How great only omniscience can say. You may speak of a chain, or if you please, a net. An analogy is of little aid. Each cause brings about future events. Without each the future would not be the same. Each is proximate in the sense it is essential. But that is not what we mean by the word. Nor on the other hand do we mean sole cause. There is no such thing.

Should analogy be thought helpful, however, I prefer that of a stream. The spring, starting on its journey, is joined by tributary after tributary. The river, reaching the ocean, comes from a hundred sources. No man may say whence any drop of water is derived. Yet for a time distinction may be possible. Into the clear creek, brown swamp water flows from the left. Later, from the right comes water stained by its clay bed. The three may remain for a space, sharply divided. But at last, inevitably no trace of separation remains. They are so commingled that all distinction is lost.

As we have said, we cannot trace the effect of an act to the end, if end there is. Again, however, we may trace it part of the way. A murder at Sarajevo may be the necessary antecedent to an assassination in London twenty years hence. An overturned lantern may burn all Chicago. We may follow the fire from the shed to the last building. We rightly say the fire started by the lantern caused its destruction.

A cause, but not the proximate cause. What we do mean by the word "proximate" is, that because of convenience, of public policy, of a rough sense of justice, the law arbitrarily declines to trace a series of events beyond a certain point. This is not logic, it is practical politics. Take our rule as to fires. Sparks from my burning haystack set on fire my house and my neighbor's. I may recover from a negligent railroad. He may not. Yet the wrongful act as directly harmed the one as the other. We may regret that the line was drawn just where it was, but drawn somewhere it had to be. We said the act of the railroad was not the proximate cause of our neighbor's fire. Cause it surely was. The words we used were simply indicative of our notions of public policy. Other courts think differently. But somewhere they reach the point where they cannot say the stream comes from any one source.

Take the illustration given in an unpublished manuscript by a distinguished and helpful writer on the law of torts. A chauffeur negligently collides with another car which is filled with dynamite, although he could not know it. An explosion follows. A, walking on the sidewalk nearby, is killed. B, sitting in a window of a building opposite, is cut by flying glass. C, likewise sitting in a window a block away, is similarly injured. And a further illustration. A nursemaid, ten blocks away, startled by the noise, involuntarily drops a baby from her arms to the walk. We are told that C may not recover while A may. As to B it is a question for court or jury. We will all agree that the baby might not. Because, we are again told, the chauffeur had no reason to believe his conduct involved any risk of injuring either C or the baby. As to them he was not negligent.

But the chauffeur, being negligent in risking the collision, his belief that the scope of the harm he might do would be limited is immaterial. His act unreasonably jeopardized the safety of any one who might be affected by it. C's injury and that of the baby were directly traceable to the collision. Without that, the injury would not have happened. C had the right to sit in his office, secure from such dangers. The baby was entitled to use the sidewalk with reasonable safety.

The true theory is, it seems to me, that the injury to C, if in truth he is to be denied recovery, and the injury to the baby is that their several injuries were not the proximate result of the negligence. And here not what the chauffeur had reason to believe would be the result of his conduct, but what the prudent would foresee, may have a bearing. May have some bearing, for the problem of proximate cause is not to be solved by any one consideration.

It is all a question of expediency. There are no fixed rules to govern our judgment. There are simply matters of which we may take account. We have in a somewhat different connection spoken of "the stream of events." We have asked whether that stream was deflected—whether it was forced into new and unexpected channels. This is rather rhetoric than law. There is in truth little to guide us other than common sense.

There are some hints that may help us. The proximate cause, involved as it may be with many other causes, must be, at the least, something without which the event would not happen. The court must ask itself whether there was a natural and continuous sequence between cause and effect. Was the one a substantial factor in producing the other? Was there a direct connection between them, without too many intervening causes? Is the effect of cause on result not too attenuated? Is the cause likely, in the usual judgment of mankind, to produce the result? Or by the exercise of prudent foresight could the result be foreseen? Is the result too remote from the cause?—and here we consider remoteness in time and space, where we passed upon the construction of a contract—but something was also said on this subject. Clearly we must so consider, for the greater the distance either in time or space, the more surely do other causes intervene to affect the result. When a lantern is overturned the firing of a shed is a fairly direct consequence. Many things contribute to the spread of the conflagration—the force of the wind, the direction and width of the streets, the character of intervening structures, other factors. We draw an uncertain and wavering line, but draw it we must as best we can.

Once again, it is all a question of fair judgment, always keeping in mind the fact that we endeavor to make a rule in each case that will be practical and in keeping with the general understanding of mankind.

Here another question must be answered. In the case supposed it is said, and said correctly, that the chauffeur is liable for the direct effect of the explosion although he had no reason to suppose it would follow a collision. "The fact that the injury occurred in a different manner than that which might have been expected does not prevent the chauffeur's negligence from being in law the cause of the injury." But the natural results of a negligent act—the results which a prudent man would or should foresee—do have a bearing upon the decision as to proximate cause. We have said so repeatedly. What should be foreseen? No human foresight would suggest that a collision itself might injure one a block away. On the contrary, given an explosion, such a possibility might be reasonably expected. I think the direct connection, the foresight of which the courts speak, assumes prevision of the explosion, for the immediate results of which, at least, the chauffeur is responsible.

It may be said this is unjust. Why? In fairness he should make good every injury flowing from his negligence. Not because of tenderness toward him we say he need not answer for all that follows his wrong. We look back to the catastrophe, the fire kindled by the spark, or the explosion. We trace the consequences—not indefinitely, but to a certain point. And to aid us in fixing that point we ask what might ordinarily be expected to follow the fire or the explosion.

This last suggestion is the factor which must determine the case before us. The act upon which defendant's liability rests is knocking an apparently harmless package onto the platform. The act was negligent. For its proximate consequences the defendant is liable. If its contents were broken, to the owner; if it fell upon and crushed a passenger's

foot, then to him. If it exploded and injured one in the immediate vicinity, to him also as to A in the illustration. Mrs. Palsgraf was standing some distance away. How far cannot be told from the record—apparently twenty-five or thirty feet. Perhaps less. Except for the explosion, she would not have been injured. We are told by the appellant in his brief "it cannot be denied that the explosion was the direct cause of the plaintiff's injuries." So it was a substantial factor in producing the result—there was here a natural and continuous sequence—direct connection. The only intervening cause was that instead of blowing her to the ground the concussion smashed the weighing machine which in turn fell upon her. There was no remoteness in time, little in space. And surely, given such an explosion as here it needed no great foresight to predict that the natural result would be to injure one on the platform at no greater distance from its scene than was the plaintiff. Just how no one might be able to predict. Whether by flying fragments, by broken glass, by wreckage of machines or structures no one could say. But injury in some form was most probable.

Under these circumstances I cannot say as a matter of law that the plaintiff's injuries were not the proximate result of the negligence. That is all we have before us. The court refused to so charge. No request was made to submit the matter to the jury as a question of fact, even would that have been proper upon the record before us.

The judgment appealed from should be affirmed, with costs.

Judgment reversed, etc.

REVIEW AND DISCUSSION QUESTIONS

1. Explain Cardozo's position. What does he mean by the statement, "Proof of negligence in the air, so to speak, will not do"? Why does he think no wrong has been done to Mrs. Palsgraf?
2. Why does Cardozo maintain that deciding in favor of Mrs. Palsgraf would involve us in a "maze of contradictions"? Do you agree?
3. Cardozo contends that "the law of causation, remote or proximate, is thus foreign to the case before us." Is it true that Cardozo's decision ignores the issue of causation altogether?
4. According to Andrews, what is "proximate" cause, and how is it different from the factual issue of what caused the injury? How is proximate cause to be determined?
5. What is Judge Andrews's position, and how does it compare with that of Cardozo? Who has the better argument, and why?
6. How is the issue of foreseeability relevant to the case, according to each judge?

Loss, Agency, and Responsibility for Outcomes

Stephen R. Perry

Corrective justice is widely thought to lie at the heart of tort law, but legal theorists frequently differ in their understanding of the concept. After explaining and criticizing the annulment conception of corrective justice, Stephen R. Perry goes on to defend what he calls the comparative conception. Influenced by philosopher Bernard Williams's discussion of "agent regret," Perry connects this idea to the notion of "outcome-responsibility." He concludes that outcome responsibility is the foundation of tort liability and the basis of whatever compensation an injurer may owe his or her victim. Determining when compensation is owed, however, is an open-ended affair, which involves weighing and comparing various factors.

INTRODUCTION

It is often claimed that the principle of corrective justice is, or ought to be, the theoretical mainspring of tort law. But this claim cannot be properly evaluated unless we know what corrective justice is. The idea that it is centrally concerned with the payment of compensation for certain losses that one person's conduct has caused another is generally accepted. It is also accepted, at least by the writers on corrective justice I shall discuss in this paper, that an important element of corrective justice is a notion of wrongfulness. Beyond this limited common ground, however, there exists deep disagreement.

THE ANNULMENT CONCEPTION

. . . This conception takes the point of corrective justice to be the elimination, rectification, or annulment of wrongful gains and losses. . . . The proposition that wrongful losses should be rectified or annulled Coleman calls the *annulment thesis*.[1] He characterizes as "wrongful" those losses that result from either wrongful conduct or the invasion of a right. Conduct is wrongful if it does not comply with relevant norms of conduct. Wrongful conduct that causes harm might or might not involve the invasion of a right. If it does, the right is said to have been *violated*. But innocent or justifiable conduct can also give rise to a rights invasion, in which case the right is said to have been *infringed*. Coleman apparently regards infringement cases as coextensive with necessity cases, in which one person takes or destroys the property of another in order to save her own life. On Coleman's analysis, the conduct of the person in the necessitous state is justified, but even so it infringes the right of the property-owner. The loss the latter suffers is thus a wrongful loss, and it falls within the purview of corrective justice. One reason why Coleman thought that wrongfulness is properly predicated, so far as corrective justice is concerned, of the victim's loss rather than the injurer's conduct was his view that in cases such as these the victim has a claim to repair even though the injurer did not act wrongfully.

Wrongful losses are thus caused, according to Coleman, by conduct that either is wrongful or invades a right. In both cases the loss is the result of an exercise of human agency, and Coleman expressly insisted that a loss could only be subject to the principle of corrective justice if it possessed this feature. The annulment conception of corrective justice says that wrongful losses should be eliminated or annulled. But the responsibility of annulment does not fall to specific persons, and in particular it does not fall to the injurer (or injurers) who caused the loss. The annulment conception does not, in other words, give rise to agent-relative reasons for action. It requires, rather, that a certain state of affairs be brought about through a reallocation of material resources within the relevant community. There might be more than one way of achieving this state of affairs, and the responsibility for doing so is in some sense a general one. . . . It is the state that would ultimately be charged with deciding how wrongful losses are to be annulled. . . .

So far as the distinction between wrongful and other kinds of losses is concerned, Coleman suggested that a person who had suffered a wrongful loss would have a *right* or a claim in justice to recover it, whereas someone who had suffered an identical loss that was not wrongful in his sense would not; the community might have moral reasons to compensate the latter victim, but no obligation or reason of justice to do so. As we have seen, however, the moral right of the former victim would ultimately be held, not against the person who had acted wrongfully or invaded her right, but against the state or community at large. Such a right would presumably be justified by the fact of loss rather than by the manner in which the loss came about, so there is no obvious reason why the victim who suffered a materially identical but non-wrongful loss should not have a similar right. Given the assumption that the responsibility for annulling losses rests with the state, the claim that there is a morally stronger case for rectifying wrongful losses than non-wrongful ones looks arbitrary.

. . . The point of a general or collective responsibility to annul losses can only be to provide compensation for the setback to well-being that a loss represents. . . . There is no justification for determining which losses are to be compensated by looking to whether they were the result of fault or, more generally, of human agency. . . .

The preceding discussion suggests that the most basic normative concept underlying the annulment thesis is not loss but rather need, which is one of the concepts that ultimately informs distributive justice at the political level. . . .

The annulment thesis goes wrong in assuming that a principle of compensation can only generate agent-neutral [reasons that do not refer to the person who has it, like reducing suffering in the world] or political reasons for action. As we shall see, however, there can also be agent-relative [reasons that refer to the individual, like something being in an individual's own interest] reasons for compensating persons who have suffered losses, and it is reasons of this kind to which corrective justice properly conceived gives rise. The justification for that principle is more complex than the argument supporting the annulment thesis, but it must still take the independent normative significance of human well-being as one of its starting points. . . .

THE COMPARATIVE CONCEPTION

A Preliminary Difficulty

As was noted in the introduction, it is common ground that corrective justice is centrally concerned with the payment of compensation for certain losses that one person's conduct has caused another. Traditionally it has also been accepted that the victim's right to repair is correlative of an obligation to compensate owed by the injurer and the injurer alone. The annulment thesis rejected correlativity and agent-relativity, and in consequence it collapsed into a theory of general distributive justice. It is of course possible that there simply is no independent principle of corrective justice, but if there is, the obligations it yields must be both limited to injurers and correlative of victims' rights. What we are seeking, then, is a justification for corrective justice that incorporates these constraints. . . .

The annulment conception begins with the reasonable premise that an interference with the well-being of one person—a loss—can affect the reasons for action of others; but it cannot limit the scope of the resulting principle to losses caused by human agency let alone insist, as in fact it does not try to do, that the relevant reasons apply to injurers alone. . . .

It might seem plausible to think that the justification we seek could . . . be something like this. Engaging in wrongful conduct has various normative consequences for the wrongdoer. The interference with individual well-being that loss represents is capable of affecting other persons' reasons for action. Where *A*'s wrongdoing causes loss to *B*, at least one of the normative consequences for *A* is that he has an obligation (i.e., a particular type of reason for action) to compensate *B* for the loss. This is because it is morally preferable that, as between a wrongdoer and an innocent victim, the former rather than the latter should suffer the interference with well-being that either the original loss or the payment of compensation entails. . . .

The justification for correlative rights and obligations of repair that we seek . . . requires that we establish responsibility, not just for wrongdoing, but for the particular loss caused. . . . I shall suggest that a more complex account of agency, in which the outcomes of actions play a role in the moral assessment of those actions and in the determination of subsequent reasons for action, supplies the missing element.

Agent-Regret, Outcome-Responsibility and Compensation

Bernard Williams has argued, persuasively and with much insight, that the assessment and justification of action is in some contexts "essentially retrospective,"[2] in the sense that it is concerned first and foremost with the outcome of action rather than with its motivation. He suggests that retrospective justification of this kind sometimes depends on the outcome of an action in a way that cannot be captured by rules for the antecedent guidance of conduct, and thus depends in part on luck. If the agent's action turns out badly, either for the agent herself or for someone else, she will ordinarily feel what Williams terms *agent-regret.* This is regret directed not just at the outcome of the action but at the action itself viewed retrospectively, in light of the outcome.

Agent-regret is not, according to Williams, restricted to cases of voluntary agency. By this he

means that the sentiment is appropriately felt not just with respect to actions giving rise to outcomes that are related to ends intentionally aimed at, but in other cases as well. Nor is it restricted to outcomes that primarily affect only the agent. In cases in which the outcome has adverse effects for someone else, agent-regret takes on a moral dimension. Williams' example is of a driver who accidentally and without fault runs over a child. The driver will, he says, experience agent-regret where a spectator would not, at least unless the spectator has the "agent's thought" that he himself might have prevented the injury. The fact that the driver has occasion to experience agent-regret is in part a matter of moral luck, since the occurrence of the accident is partly due to factors beyond his control. . . .

A spectator who witnessed the running over of the child might regret the incident, in the sense of wishing that things had been otherwise, but he would not have a wish that *he* had acted otherwise. The agent-regret of the driver and the regret of the spectator differ not only in the content of the sentiment felt, but also in the way their respective sentiments are expressed: "The lorry-driver may act in some way which he hopes will constitute or at least symbolize some kind of recompense or restitution and this will be an expression of agent-regret." This is a crucially important point, because it suggests that agent-regret is not just a sentiment but a truly normative notion that can directly affect present reasons for action. We have moved from the retrospective evaluation of one's past actions to a conception of responsibility of some kind for the outcomes of those actions. . . . Let me label such conceptions of responsibility, both subjective and objective, as "outcome-responsibility." It is important to make clear that outcome-responsibility does not necessarily carry with it an obligation to compensate, or even a non-peremptory reason to do so. It is a more general normative notion that may, for example, simply give rise to reasons to obtain assistance, or to express one's regret to the person harmed. But while outcome-responsibility may not be a sufficient condition for recognizing an obligation of repair it is, I wish to suggest, a necessary condition: it is the *basis* of an obligation to compensate. . . .

What is the basis of agent-regret itself, and why does it ground a conception of responsibility for outcomes? Why should we think that even in cases of culpable harm there is responsibility for the particular outcome, as opposed to responsibility simply for behaving badly? . . . As agents we act with the intention of producing an effect of some sort under a certain description, and our sense of agency is such that when we succeed we regard the effect, in an important and fundamental sense, as ours: by exercising our powers of voluntary agency we have made a difference in the world, and we identify with the result. In cases of harm-producing actions involving wrongful intent it is appropriate that we regard the action with agent-regret, which in such a case means wishing that we had acted otherwise so as not to have produced the harmful outcome. But we cannot change the past, so agent-regret properly takes on a normative aspect that may affect our present and future reasons for action vis-à-vis the outcome in question. It is transformed, in other words, into a conception of outcome-responsibility. . . .

Similar considerations apply to at least some unintended consequences. As persons capable of self-reflective thought we are aware that we do not have complete control over either our own character and identities as agents, or the physical environment in which we act. Things may, for a variety of reasons, turn out differently from what we expect. . . . But the fact that things turn out unexpectedly is not a sufficient reason to disassociate ourselves from the outcomes we produce. As in the case of intended outcomes, we are aware that our exercise of agency has made a difference in the world, and given our knowledge of the circumstances under which we necessarily act (i.e., our lack of complete control over our own natures and our environment) it may be that we should likewise identify with at least some outcomes that were unintended. If so, they too will be ours, and if they involve adverse effects for others we should similarly experience agent-regret and accept the possibility of its transformation into a normative conception of outcome-responsibility.

When should we identify with unintended consequences in this way? Although it will not be possible to justify the following answer here, I wish to suggest that we ordinarily do and should

so identify *at least* in circumstances which are of such a character that we would generally be capable, in situations of that kind, of acting so as to avoid the harm, even if we were not capable of so acting on the particular occasion due to, in the usual case, a failure of foresight. Because of the close connection between avoidability and foreseeability, the resulting conception of responsibility is an essentially epistemic one: it coincides for the most part with the notion of reasonable foreseeability, which is the most important determinant of a duty of care in negligence law. This conception of outcome-responsibility is not equivalent to fault, however, since fault is concerned with conduct that would constitute *breach* of a duty of care. Outcome-responsibility thus understood is objective in two senses. First, it does not depend on the subjective capacities of the particular agent to have acted otherwise on the particular occasion . . . and second, it is appropriately attributed to agents even if they did not happen on the occasion in question to feel the sentiment of agent-regret. The concept of outcome-responsibility is rooted in the sense of agent-regret that persons normally feel with respect to certain of their actions that produced harmful but avoidable outcomes, but it is not simply to be identified with that sentiment.

If the points urged in the preceding paragraph are correct, the heart of outcome-responsibility is a notion of avoidability. Avoidability is not the same as fault, since there is no necessary connotation that the agent *should* have acted otherwise. But a person who experiences agent-regret will nonetheless wish that she had acted otherwise, or at least that she had not had to act as she did, so that the harm in question would not have occurred. But the past cannot be undone, and hence the normative focus shifts to the present and future. The agent's reasons for action are affected, but this does not necessarily carry with it an obligation to compensate; the agent may have reasons only to obtain assistance, express regret, or make what Williams calls symbolic recompense. The possibility is at least raised, however, that she should pay full compensation. . . . This sets the stage for a comparative inquiry of the following kind: which of these two [victim or injurer] should most appropriately suffer the interference with well-being that *either* the original loss *or* the payment of compensation necessarily entails?

. . . In general, the comparative inquiry focuses primarily but not exclusively on two factors: first, on the nature of the injurer's conduct in causing the damage, and in particular on whether her conduct exhibited fault of some kind, and second, on the nature of the victim's detrimentally affected interest. These two factors are not independent of one another. . . .

Consider . . . the difference between the law of negligence and the law of nuisance: the former determines liability by reference to a certain level of permissible risk, whereas the latter determines liability by reference to a certain level of permissible *harm.* This means that the interest protected by nuisance, namely, the use and enjoyment of real property, can permissibly be harmed intentionally, at least up to a point, although if the motivation is malice the law may call the interfering action a nuisance anyway and impose legal liability. Another point to be noted is that the comparative inquiry, and consequently tort law, may impose an obligation to compensate for certain kinds of justified conduct that would nonetheless be wrongful in the absence of a specific justifying factor. Necessity cases such as *Vincent v. Lake Erie Transportation Co.* are of this type, for example. The defendant's action in intentionally damaging the plaintiff's property is justified by the need to counter some serious threat to the former's life or health. The action is nonetheless in a certain sense fault-like, since its intentional nature would make it wrongful were it not for the situation of necessity. This fault-like aspect is what tips the balance and justifies, on the comparative approach, the payment of compensation. . . .

As the previous discussion makes clear, there is no single criterion of distribution, such as a particular conception of fault, that can be applied in a uniform way in all cases. Rather the inquiry is an open-ended one: it is possible to take many different factors into consideration, and there is no preestablished list of what these are. . . . The victim is included because in the nature of things he will have to continue to bear the loss he has experienced if no one else has a reason to take it on. The injurer is included because she is outcome-responsible for the loss, and hence potentially has a

reason to shoulder it herself. The comparative inquiry is simply a process of determining whether, all things considered, the injurer really does have such a reason, in the form of an obligation to pay compensation. . . .

The discussion of the preceding three paragraphs suggests that the comparative inquiry called for by the comparative conception of corrective justice is a fairly open-ended affair, which can take a number of different kinds of consideration into account. This suggests in turn that corrective justice is less a matter of fixed substantive principles than it is a framework for certain kinds of moral argument.

NOTES

1. See, e.g., Coleman, "Tort Law and the Demands of Corrective Justice" (1992) 67 *Indiana L.J.* 349, 357; Coleman, "Property, Wrongfulness and the Duty to Compensate" (1987) 63 *Chi.-Kent L. Rev* 451; Coleman, "Moral Theories of Torts: Their Scope and Limits: Part II" (1983) 2 *Law & Phil.* 5; Coleman, "Corrective Justice and Wrongful Gain" (1982) 11 *J. of Legal Stud.* 421.
2. Williams, "Moral Luck," in his *Moral Luck* (New York: Cambridge University Press, 1981) 20, at 24. There is also a very useful and pertinent discussion of moral luck in Nagel, "Moral Luck," in his *Mortal Questions* (1979) 24.

REVIEW AND DISCUSSION QUESTIONS

1. What is the annulment conception of corrective justice? On what grounds does Perry criticize it? Do you agree that the annulment conception is an unsatisfactory basis on which to ground tort liability? Explain why or why not.
2. Explain what Williams has to say about luck, the outcome of our actions, and "agent regret."
3. According to Perry, what is the basis of agent regret, and how does it ground a conception of responsibility for the outcomes of our actions? Why does outcome responsibility sometimes apply to the unintended consequences of our actions? What is the connection between outcome responsibility and "avoidability"?
4. Explain the nature of the "comparative inquiry" that Perry says is necessary to determine whether compensation is to be paid.
5. Is Perry's conception of corrective justice an improvement over the annulment conception? Does it provide a satisfactory underpinning for tort law?

Liability Without Causation?

Summers v. Tice

The plaintiff Charles Summers went quail hunting with Harold Tice and Ernest Simonson. When a quail was flushed, Tice and Simonson fired negligently in Summers's direction, and he was wounded in the face. Because Tice and Simonson fired similar pellets from similar guns, however, it could not be proved which of them, or which gun, had actually injured Summers. Although a basic element of negligence as a tort is that the defendant's conduct must be the cause of the plaintiff's injury, in this case the California Supreme Court upheld Summers's suit.

---

199 P.2d (1948).

Actions by Charles A. Summers against Harold W. Tice and against Ernest Simonson for negligently shooting plaintiff while hunting. From judgments for plaintiff, defendants appeal. . . .

Justice Carter delivered the opinion of the Court: Each of the two defendants appeals from a judgment against them in an action for personal injuries. Pursuant to stipulation the appeals have been consolidated.

Plaintiff's action was against both defendants for an injury to his right eye and face as the result of being struck by bird shot discharged from a shotgun. The case was tried by the court without a jury and the court found that on November 20, 1945, plaintiff and the two defendants were hunting quail on the open range. Each of the defendants was armed with a 12 gauge shotgun loaded with shells containing 7 1/2 size shot. Prior to going hunting plaintiff discussed the hunting procedure with defendants, indicating that they were to exercise care when shooting and to "keep in line." In the course of hunting, plaintiff proceeded up a hill, thus placing the hunters at the points of a triangle. The view of defendants with reference to plaintiff was unobstructed and they knew his location. Defendant Tice flushed a quail which rose in flight to a ten foot elevation and flew between plaintiff and defendants. Both defendants shot at the quail, shooting in plaintiff's direction. At that time defendants were 75 yards from plaintiff. One shot struck plaintiff in his eye and another in his upper lip. Finally it was found by the court that as the direct result of the shooting by defendants the shots struck plaintiff as above mentioned and that defendants were negligent in so shooting and plaintiff was not contributorily negligent.

. . . First, on the subject of negligence, defendant Simonson contends that the evidence is insufficient to sustain the finding on that score, but he does not point out wherein it is lacking. There is evidence that both defendants, at about the same time or one immediately after the other, shot at a quail and in so doing shot toward plaintiff who was uphill from them, and that they knew his location. That is sufficient from which the trial court could conclude that they acted with respect to plaintiff other than as persons of ordinary prudence. . . .

Defendant Tice states in his opening brief, "we have decided not to argue the insufficiency of negligence on the part of defendant Tice." It is true he states in his answer to plaintiff's petition for a hearing in this court that he did not concede this point but he does not argue it. Nothing more need be said on the subject.

. . . Defendant Simonson urges that plaintiff was guilty of contributory negligence and assumed the risk as a matter of law. He cites no authority for the proposition that by going on a hunting party the various hunters assume the risk of negligence on the part of their companions. Such a tenet is not reasonable. It is true that plaintiff suggested that they all "stay in line," presumably abreast, while hunting, and he went uphill at somewhat of a right angle to the hunting line, but he also cautioned that they use care, and defendants knew plaintiff's position. We hold, therefore, that the trial court was justified in finding that he did not assume the risk or act other than as a person of ordinary prudence under the circumstances. . . .

The problem presented in this case is whether the judgment against both defendants may stand. It is argued by defendants that they are not joint tort feasors, and thus jointly and severally liable, as they were not acting in concert, and that there is not sufficient evidence to show which defendant was guilty of the negligence which caused the injuries—the shooting by Tice or that by Simonson. Tice argues that there is evidence to show that the shot which struck plaintiff came from Simonson's gun because of admissions allegedly made by him to third persons and no evidence that they came from his gun. Further in connection with the latter contention, the court failed to find on plaintiff's allegation in his complaint that he did not know which one was at fault—did not find which defendant was guilty of the negligence which caused the injuries to plaintiff.

. . . Considering the last argument first, we believe it is clear that the court sufficiently found on the issue that defendants were jointly liable and that thus the negligence of both was the cause of the injury or to that legal effect. It found that both defendants were negligent and "That as a direct and proximate result of the shots fired by *defendants, and each of them*, a birdshot pellet was caused to and did lodge in plaintiff's right eye and that another birdshot pellet was caused to and did lodge in plaintiff's upper lip." In so doing the court evidently did not give credence to the admissions

of Simonson to third persons that he fired the shots, which it was justified in doing. It thus determined that the negligence of both defendants was the legal cause of the injury—or that both were responsible. Implicit in such finding is the assumption that the court was unable to ascertain whether the shots were from the gun of one defendant or the other or one shot from each of them. The one shot that entered plaintiff's eye was the major factor in assessing damages and that shot could not have come from the gun of both defendants. It was from one or the other only.

It has been held that where a group of persons are on a hunting party, or otherwise engaged in the use of firearms, and two of them are negligent in firing in the direction of a third person who is injured thereby, both of those so firing are liable for the injury suffered by the third person, although the negligence of only one of them could have caused the injury. . . . *Oliver v. Miles*, Miss., 110 So. 666, 50 A.L.R. 357. . . . The same rule has been applied in criminal cases . . . and both drivers have been held liable for the negligence of one where they engaged in a racing contest causing an injury to a third person. . . . These cases speak of the action of defendants as being in concert as the ground of decision, yet it would seem they are straining that concept and the more reasonable basis appears in *Oliver v. Miles*, supra. There two persons were hunting together. Both shot at some partridges and in so doing shot across the highway injuring plaintiff who was traveling on it. The court stated that they were acting in concert and thus both were liable. The court then stated. . . : "We think that . . . each is liable for the resulting injury to the boy, although no one can say definitely who actually shot him. *To hold otherwise would be to exonerate both from liability, although each was negligent, and the injury resulted from such negligence.* . . . "

. . . When we consider the relative position of the parties and the results that would flow if plaintiff was required to pin the injury on one of the defendants only, a requirement that the burden of proof on that subject be shifted to defendants becomes manifest. They are both wrongdoers—both negligent toward plaintiff. They brought about a situation where the negligence of one of them injured the plaintiff; hence it should rest with them each to absolve himself if he can. The injured party has been placed by defendants in the unfair position of pointing to which defendant caused the harm. If one can escape the other may also and plaintiff is remediless. Ordinarily defendants are in a far better position to offer evidence to determine which one caused the injury. This reasoning has recently found favor in this Court. In a quite analogous situation this Court held that a patient injured while unconscious on an operating table in a hospital could hold all or any of the persons who had any connection with the operation even though he could not select the particular acts by the particular person which led to his disability. *Ybarra v. Spangard* . . . 154 P.2d 687. . . . There the Court was considering whether the patient could avail himself of res ipsa loquitur, rather than where the burden of proof lay, yet the effect of the decision is that plaintiff has made out a case when he has produced evidence which gives rise to an inference of negligence which was the proximate cause of the injury. It is up to defendants to explain the cause of the injury. . . .

The foregoing discussion disposes of the authorities cited by defendants . . . , stating the general rule that one defendant is not liable for the independent tort of the other defendant, or that ordinarily the plaintiff must show a causal connection between the negligence and the injury. There was an entire lack of such connection in the Hernandez case and there were not several negligent defendants, one of whom must have caused the injury. . . .

In view of the foregoing discussion it is apparent that defendants in cases like the present one may be treated as liable on the same basis as joint tort feasors, and hence the last cited cases are distinguishable inasmuch as they involve independent tort feasors.

. . . It should be pointed out that the same reasons of policy and justice shift the burden to each of [the] defendants to absolve himself if he can—relieving the wronged person of the duty of apportioning the injury to a particular defendant, apply here where we are concerned with whether plaintiff is required to supply evidence for the apportionment of damages. If defendants are independent tort feasors and thus each liable for the damage caused by him alone, and, at least, where the matter of apportionment is incapable of proof, the innocent wronged party should not be deprived of his right to redress. The wrongdoers should be left to work out

between themselves any apportionment. . . . Some of the cited cases refer to the difficulty of apportioning the burden of damages between the independent tort feasors, and say that where factually a correct division cannot be made, the trier of fact may make it the best it can, which would be more or less a guess, stressing the factor that the wrongdoers are not in a position to complain of uncertainty.

. . . It is urged that plaintiff now has changed the theory of his case in claiming a concert of action; that he did not plead or prove such concert. From what has been said it is clear that there has been no change in theory. The joint liability, as well as the lack of knowledge as to which defendant was liable, was pleaded and the proof developed the case under either theory. We have seen that for the reasons of policy discussed herein, the case is based upon the legal proposition that, under the circumstances here presented, each defendant is liable for the whole damage whether they are deemed to be acting in concert or independently.

The judgment is affirmed.

REVIEW AND DISCUSSION QUESTIONS

1. What is the court's reasoning in upholding Summers's suit against Tice? Do you agree that Tice and Simonson were jointly liable?
2. Would it have been unfair to Summers to deny him compensation because he could not prove who shot him? The Court explicitly places the burden of proof on Tice to show that he was not the cause of Summers's injury, rather than on Summers to show that he was injured by Tice. Is this reasonable and fair?
3. One defendant—either Tice or Simonson—caused Summers no harm whatsoever. Is it just that this defendant be required to compensate Summers for an injury for which he (the defendant) is not responsible?
4. In this case, both Tice and Simonson shot in Summers's direction. Suppose only one of them had done so, but it could not be proved which one. Should Summers still win his suit?
5. When is it fair, and when is it unfair, to impose tort liability without proof of causation?

Legal Causation and the Desert Traveler

Amy Shapiro

The question of whether an act caused a particular outcome is an important one in the law. Particularly in tort law, liability depends on whether what a person did caused harm. If no harm is caused by a person's negligence, there is no legal liability, but if someone is injured, then the person who caused the injury must make compensation. However, it can be difficult to discern whether someone caused something, especially in complex factual situations. Because of difficulties with the traditional test of causation (the "but-for" test), courts may appear to impose liability for harm without requiring a causal link to the defendant's actions.

In this essay, Amy Shapiro argues that situations where courts might appear to impose liability without causation actually do involve causation, though in different forms from the traditional,but-for test of legal causation. The but-for test involves only one form of ordinary, non-legal causal thinking, namely determining causation for the purpose of explaining something. In everyday life, people also use causation to make predictions and warnings and to assign responsibility. The law uses causation in all three of these everyday senses. Shapiro applies the different senses of causation to a classic puzzle in legal causation, the hypothetical case of the Desert Traveler. Amy Shapiro is a practicing lawyer who also teaches law at Cornell Law School.

© 2004, Amy Shapiro. All rights reserved.

A man is about to travel on foot alone into the desert. He has two enemies. Unfortunately for him, each of them finds out about the planned trip. The night before, while the Traveler is sleeping, Enemy No. 1 quietly sneaks up and poisons his water supply. A little later, Enemy No. 2—who knows nothing about Enemy No. 1—drills holes in the water carrier. All the poisoned water drains out into the sand.

In the morning, the Traveler packs up and leaves. Because the gear is so heavy, it is not obvious that the water carrier is empty. Somewhere in the desert, the Traveler dies of thirst. The question is: who caused his death? Enemy No. 1, the Poisoner? Enemy No. 2, the Driller? Both? Neither?

This classic hypothetical of the Desert Traveler[1] presents an unusual situation. But it illustrates some of the knottiest problems in legal causation. Causation is important in the law, especially tort law, because it is generally viewed as the crucial link between what a defendant did and what harm the plaintiff suffered as a result. This link is what justifies a requirement by the court that the defendant make compensation. Just what is the proper test of legal causation, however, is a difficult and contentious question.

Discussions of causation tend to focus so much on the traditional "but-for" test (to be discussed later) that other approaches to figuring out the link between an act and an outcome can be overlooked. Indeed, compared to the but-for test, other tests may seem to excuse the requirement of causation altogether and impose liability on some other basis. What perhaps has not been fully appreciated is that different methods of seeking to link act and outcome *are* tests of causation, though they approach it in different ways. In what follows, I will argue that the law's various approaches to causation fall into three basic groups, and these track the three basic functions of causation in everyday life: explaining events, making predictions and assigning responsibility.[2]

Therefore, as I will discuss, the traditional "but-for" test is really an attempt to find causal explanation. It is backward looking. It focuses on the individual event in question and the particular link between act and outcome. The problem with using causation as explanation, however, is that it can identify both too many causes and too few, as we will see in the case of the Desert Traveler. The relatively newer use of statistical correlation to serve the function of causation is based on the practice of prediction. This is forward-looking, in that it takes the point of view of the agent at the time of the action: what was predicted to happen? The problem with causation as prediction, however, is that in law liability is individual, whereas a prediction cannot precisely connect a particular past act and a particular outcome. Rather, it shows the usual, expected connection between this sort of act and this sort of outcome, producing a correlation that is probably—but not necessarily—true in the given situation. Finally, a third test of causation used in U.S. courts directs the jury to find that an event caused an outcome if the event was a "substantial factor" in bringing the outcome about. This tracks the everyday activity of using causation to assign responsibility. Here, the potential problem is that the test is undisciplined and vague, giving jurors little guidance. I will conclude that especially in hard cases like the Desert Traveler, this open-endedness turns out to be a virtue, not a vice. Untidy though it may be, the substantial factor test is realistic given unsettled debates over the proper tests of causation in the law; it allows jurors to employ the test best suited to the case at hand, and it focuses the causal question where it belongs: Who is responsible?

THREE TYPES OF CAUSAL INQUIRIES

Explanation

The first kind of causal inquiry is historical. It is done for the purpose of explaining something that has happened: the drought caused the forest fire; entangling European alliances caused World War I. Explanation involves looking at something that occurred in the past and reasoning backwards from the result to determine what brought it about.

A satisfactory causal explanation will not discuss all possible causes. Instead, it must identify some things as having made a difference and relegate others to the background, depending on the purpose of the explanation. For example, what caused Johnny's birthday cake to fall? Looking at

the same facts, a baker might say that there was no baking soda, a scientist that certain chemical reactions failed and a psychologist that Johnny's aunt was annoyed at him when she made the cake. Each explanation is complete in its own terms. The psychological story does not need the chemical explanation to complete it and vice versa. The baker's diagnosis of "no baking soda" stands on its own without knowing why the soda was left out or how baking soda works chemically.

Prediction

Many causal explanations are really intended to be the basis of a prediction of the future. In asking whether x caused y, what we really want to know is whether the next time x happens, will y occur, even if other factors (z) are different? Will the dog that bit you when you came to the door bite again even if you are not wearing your Postal Service uniform? Making predictions and giving warnings (which are also predictions) is perhaps the prime role of causation in science, technology and everyday life. This includes both predictions about what will follow from an existing or assumed state of affairs ("the pit bull will bite the mailman") and "recipes" telling how to achieve a desired future state (to lose weight, eat less and exercise more; to bake a cake, use these amounts of flour, sugar, butter and baking soda). These are causal statements because they predict that doing something (delivering mail, eating less, mixing ingredients) will cause a result (dog bite, weight loss, cake).

In predictions, the goal is to describe a set of circumstances sufficient to produce a particular outcome, rather than to focus on any particular factor. It would be absurd to ask whether flour "causes" a cake. And, there is no need to rule out other possible causes of the same outcome. It is perfectly reasonable to conclude that diet and exercise cause weight loss without also noting that amputation of a limb causes it too.

Predictions rely on extrapolating from what has usually happened in the past to what will likely happen in the future. For instance, imagine that balls flew out of a baseball stadium into a neighboring street three times last season. Considering that balls were hit thousands of times, it is highly unlikely that any one will land in the street. Therefore, we would predict that it is safe to play ball at that stadium even though in the unusual case, a ball could injure someone outside the park. This injury will not be *predictable*, but once it has happened, it can be *explained*—the ball flew in a particular direction, hitting the passerby. If this happens, the prediction did not fail. All it said was that in the ordinary case, the ball would not go into the street. That statement is true even if the non-ordinary case occurs.[3]

Attribution of Responsibility

The third purpose of causal thinking, besides explanation and prediction, is to decide whether an outcome flowed from a person's actions such that she is responsible for the outcome, good or bad.[4] Like predicting and explaining, this is a common, everyday activity. Was Mike responsible for breaking up Martha's marriage (or was it already on the skids)? Was Fred responsible for the auto accident for not taking away Bob's keys when he left the bar (or is drunken Bob solely to blame)? On a happier note, was Lucia responsible for her virtuoso recital (or should the kudos go to her violin teacher for inspiring a talented but lazy student to practice)?

Attributing responsibility is serious business since with responsibility generally come consequences. Legal responsibility often means that a person must pay compensation, lose property or suffer a criminal penalty. But other kinds of responsibility also matter. Being held morally responsible for something by your community or by an authority you respect (friend, clergy or, especially, your own conscience) can have lasting repercussions.

It is notable that responsibility is an on-off switch. A person either is or is not responsible for an outcome. Saying that he is "partly responsible" is still a determination of responsibility, namely shared responsibility. But to say he "might" be responsible just means that we don't know whether he is or not. This creates a mismatch. People understand that an historical or scientific explanation may be subject to correction and that a prediction might not come true, but when causation is used to assign responsibility, the paradox arises that a factual conclusion that is often not

completely certain—namely causation—is used to justify an outcome—namely responsibility—that takes on an air of certainty.

Nevertheless, we obviously do assign causal responsibility all the time. Informal practices recognize and adjust for uncertainty. Intention may be a guide to the degree of certainty needed. For instance, if a waiter spills coffee in a customer's lap, the waiter, though he caused the burns in an explanatory sense, is not held responsible—it was an accident. However, if the waiter and customer were bitter enemies, the waiter is more likely to be held responsible unless it is clear that the spill was an unavoidable accident. We also ask for more certainty when assigning responsibility for a grave outcome (ruining a marriage vs. ruining a dinner party) or when the decision-maker's attribution will be taken seriously (the word of a boss vs. the gossip of a co-worker). The point is that the social practice of assigning responsibility is able to deal with the difficult issue of causal uncertainty.

The social practice deals with another difficult issue as well, namely the fact that people disagree about where responsibility comes from: is it somehow in the nature of things, such that causal responsibility is discovered in the same way we might discover that virus *a* rather than virus *b* caused an epidemic? Or, is it simply a social decision that someone should be blamed for a particular act, such that responsibility is *created* by the pronouncement of the appropriate authority? Interestingly, the *practice* of assigning responsibility seems to be the same no matter which theoretical position a person might take. Bob's widow might sincerely hold the second view and, indeed, have just finished writing an essay on the "Social Construction of Responsibility" when she learns that Bob died in an auto accident. Nevertheless, she might still tell Fred: "I hold you responsible for Bob's death because you allowed him to drive when he was drunk." And she would almost certainly mean to say that Fred *is* responsible for the death, and not merely that she holds that opinion. This point—that the social practice of assigning responsibility goes on independently of theoretical views of what responsibility is—will be important later in discussing causation in the law.

LEGAL CAUSATION

Though it might seem that determining causation for the purpose of assigning legal responsibility would be like determining causation for the purpose of assigning other kinds of responsibility (the third type of causation discussed above), legal doctrines of causation in fact make use of all three everyday uses of causation: explanation, prediction and assignment of responsibility.

As a threshold matter, causation is important when the law is concerned about an outcome of an act, in contrast to legal rules that penalize merely doing an act—e.g., reckless driving or driving under the influence of alcohol are legal violations whether anyone is hurt or not. Most scholarly debate concerns causation in tort law because tort liability almost always depends on outcome.[5] Merely being careless is not a tort. But if violating a duty of care causes damages to someone, then the wrongdoer is liable. Causation is what links the tortuous act with the resulting injury in order for it to be just to order the defendant to compensate the plaintiff.

Legal Causation as Explanation: Cause in Fact and Proximate Cause

Traditionally, law begins the inquiry into causation using the "but-for" test. Its everyday analogue is causal explanation; it seeks to understand what happened in the past. The but-for test directs the fact finder to imagine a world in which everything happened as it did in the case at hand, except for (i.e., *but for*) the allegedly causal act. If the outcome would have happened anyway, then the act being scrutinized is not its cause.

Every situation has a virtually infinite number of "but-for" causes. If Bob's friend hadn't invited him to the bar, for example, Bob would not have been drinking, would not have driven drunk and would not have been killed in the accident. "But for" the births of Bob, the friend and the bartender who served the last drink, the evening would not unfolded as it did, which means that the meetings of each of the sets of parents are also but-for causes. Yet, courts ignore these many but-for causes and focus on what the defendant did.

This tracks the ordinary practice of causal explanation, which is always made for a reason, not in a vacuum. The purpose of an inquiry dictates which factors will be brought to the fore and which left in the background, as with the example of the baker, psychologist and chemist each focusing on different facts to explain what ruined the cake. Since the purpose of determining causation in law is to decide whether the defendant should compensate the plaintiff, the focus is on whether the defendant's acts were the "the cause without which the event could not have occurred,"[6] notwithstanding the infinity of other but-for causes.

The "but-for" test usually works well. Take the example of a hotel guest who is assaulted by someone who entered her room through a door that was easily opened without a key. The guest sues the hotel for compensation for her injuries, since it owes guests the duty to take ordinary care for their safety. The but-for test would ask whether the assailant could have entered "but for" the broken lock. If a working lock would have kept him out, then the hotel is liable—which makes sense.[7] In other situations, however, the results of the but-for test do not comport so easily with our intuitions about the defendant's liability. The defendant's actions may be a but-for cause, but still seem too remote to justify liability. Or the "but-for" test may excuse a defendant who seemed very much involved. The court grappled with the first of these problems in *Palsgraf v. Long Island Railroad*,[8] a case in which the "but-for" test pointed the causal finger at a defendant whose actions seemed remote from the plaintiff's injury.

The facts of *Palsgraf* are extraordinary: (1) a railway conductor pushed a man onto a train, (2) the man dropped a package wrapped in newspaper, (3) the package contained fireworks, (4) the fireworks exploded, (5) the explosion toppled a heavy scale at the other end of the train platform, and (6) the falling scale injured Helen Palsgraf. As long as this causal chain may be, the conductor's pushing the man onto the train clearly forms the first link. "But for" that, the scale would not have fallen on Mrs. Palsgraf. A jury therefore found the railroad (as the conductor's employer) liable for her injuries. The railroad appealed the decision to the highest court of New York.

On appeal, the court set aside the jury's verdict. In an influential decision, Justice Cardozo held that the railroad had no legal duty to take care to avoid injuring Mrs. Palsgraf because "there was nothing in the situation to suggest to the most cautious mind that the parcel wrapped in newspaper would spread wreckage through the station." In other words, it is not negligent to fail to prevent harm that could not have been reasonably foreseen. By focusing on the duty, Justice Cardozo managed to avoid having to consider whether the conductor's action caused the injury.

Justice Andrews wrote an equally important dissenting opinion in *Palsgraf* In his view, the railroad did have a duty to avoid the accident, so that whether the conductor had caused it was the decisive question. He believed that it was reasonable for the jury to find that he had. Pushing the passenger onto the train was "a substantial factor in producing the result," which followed in "a natural and continuous sequence." Justice Andrews recognized the potentially endless reach of the "but-for" test, which he equated with causation in general. Causation is like "ripples in a pool" that extend out indefinitely. When, he asked, should the law stop following the ripples? The causal inquiry cannot answer this question. Instead, "because of convenience, of public policy, of a rough sense of justice, the law arbitrarily declines to trace a series of events beyond a certain point," he wrote. "This is not logic. It is practical politics." This conclusion is troubling. If reasons such as "practical politics" are determining causation, then how can the tort law goal of matching consequences (liability) with what a person did in the world (cause an injury) be fulfilled?

The problem is exacerbated when the "but-for" test fails to identify as a cause a defendant who nevertheless seems responsible for an injury. Does justice then require ignoring this finding of "no causation" and imposing liability without causation? Here we return to the Desert Traveler. The "but-for" test asks: Would the Traveler have died "but for" Poisoner's poisoning the water? The answer is yes, he would have died anyway because Driller emptied the water out. Therefore, Poisoner did not kill the Traveler. The but-for test then asks whether the Traveler would have died "but for" Driller's draining the water. Again, the answer is

yes. Driller spilled out poisoned water. Had he not done that, the Traveler would have died anyway, except from poisoning rather than thirst. So, the but-for test tells us that neither Poisoner nor Driller was the cause of death. But the Traveler is definitely dead, and it must be true that *somebody* killed him. The but-for test cannot be right.

Similar cases of causal over-determination can be imagined. Two fires from different sources may spread to a building at the same time,[9] or the noise from two motorcycles may simultaneously scare a horse, causing injury. Assuming that each fire or each motorcycle by itself was enough to destroy the building or scare the horse, again, the "but-for" test produces the weird result that neither is the cause. In the face of this conundrum, J. A. McLaughlin, the author of the Desert Traveler problem, concluded that a causal inquiry could not answer the question of causation in situations where there seem to be too many causes (i.e., over-determination). Anticipating the reasoning of Justice Andrews, he concluded that responsibility for killing the Traveler can only be a question of "law or policy," not fact.

The McLaughlin/Andrews line of thinking seems to leave little use for the doctrine of causation in the law. It suggests that whether to impose legal liability inevitably rests on a debate over law or policy, and not on a determination of what in fact caused the harm. Tort judgments then look like merely governmental decisions (through the court system) to make some people pay the cost of other people's injuries because that is more socially advantageous. This treats the people involved—both tortfeasors and victims—as only means to implementing social policy, rather than as responsible agents who are either required to pay compensation or entitled to receive it on the basis of responsibility for the harm actually caused. If the moral authority of law (particularly tort law) rests on its ability to hold people responsible for the natural consequences of their own actions,[10] then law's legitimacy is compromised.

Were the but-for test the only test of causation-in-fact, the conclusion that liability is only a matter of law or policy might be difficult to avoid. Problems with the but-for test, however, do not require abandoning law's reliance on factual causation in general.[11] Rather, they show why the law should not rely exclusively on causation-as-explanation. It is only one of the commonly understood senses of causation. The other everyday uses of causation can and do inform tests of legal causation. And, they can answer causal questions that the but-for test cannot handle.

Legal Causation as Prediction: Statistical Correlation and Sufficient Sets

The search for a causal explanation (the "but-for" test) looks backwards from the time of the legal proceedings. Another approach is to imaginatively step into the shoes of the defendant *before* the act in question and ask what outcome would be the predicted result. Cardozo used this technique to limit the reach of causation in *Palsgraf* by saying that the possibility that the scale would fall on Mrs. Palsgraf was not foreseeable to the conductor at the time he pushed the passenger onto the train (and therefore, the conductor had no duty to try to avoid that possible harm). Courts also use causation-as-prediction not just to limit causation, but to find it in the first place, as shown by two California Supreme Court cases.

In *Sindell v. Abbot Laboratories*,[12] a class-action products liability case, the plaintiffs were cancer victims whose mothers had taken DES (diethylstilbestrol) during pregnancy. DES was clearly connected with the cancers. However, some 200 companies had manufactured DES at the relevant time, and the lead plaintiff could not prove which one had made the DES that her mother took. Under the usual rule, this would mean she could not recover for her injuries, since plaintiffs have the burden of proof on all elements of the lawsuit, including causation. Unless the plaintiff can show the causal link between her injury and a particular defendant's actions, that defendant need not pay compensation.

The court, however, did not hold the plaintiff to this burden of proof, but shifted it to the defendants. Unless each defendant could prove that the plaintiff's mother did not take its DES pills (which would be virtually impossible), that defendant would be liable for the proportion of the plaintiff's injuries "represented by its share of the market." The court reasoned that this was fair because of

probabilities. All DES is alike. Thus, a company that manufactured one-fifth of all DES predictably would be responsible for one-fifth of all injuries. Mathematically, the court observed, the same result is produced if the company simply is held responsible in one case for one-fifth of the plaintiff's injuries.

Notably, by the time the case reached the California Supreme Court only five manufacturers were left as defendants. Although they had made a "substantial share" of all DES, there was no proof whatsoever that they manufactured the actual DES that caused the cancer in the plaintiffs. It might have been made by any of them, or by one of the 195 companies not before the court. Nevertheless, said the court, the five manufacturers would have to share all the cost of compensating the plaintiffs for their injuries.

Does this mean that the court abandoned the traditional tort law idea that liability is based on an outcome, not merely an act, and held the defendants responsible for DES injuries just because they manufactured DES? Not necessarily. Seen another way, the court did not excuse the causation requirement, but rather switched from a model of causation-as-explanation to one of causation-as-prediction. There was causation because a certain number of DES-related cancers could be predicted to follow from the defendants' business activities.

Prediction-based liability fits comfortably with the facts of *Sindell.* Predictions are based on statistical aggregates—what is likely to happen in general. The companies, patients and doctors in *Sindell* themselves treated DES manufacturing as an aggregate activity. "DES was produced from a common and mutually agreed upon formula as a fungible drug interchangeable with other brands of the same product," the court noted, and "defendants knew or should have known that if was customary for doctors to prescribe the drug by its generic rather than its brand name and that pharmacists filled prescriptions from whatever brand of the drug happened to be in stock." No one cared which factory any DES tablet came from any more than a person buying flour at the supermarket cares what farm produced the wheat. True, from the point of view of the but-for test and causation-as-explanation, only one company, the actual producer of the pills taken, injured any particular plaintiff, but it would seem wrong to insist an injured plaintiff prove who made which of billions of identical pills when that didn't matter at the time of the relevant acts. Rather, just as a manufacturer depends on probabilities in figuring out everything from materials standards to cost-effective marketing, it makes sense to think about its responsibility in statistical terms. A known probability of causing harm becomes a good reason to impose liability.

The California court also imposed liability based on a probability of causing harm in another well-known case, *Summers v. Tice.*[13] Here, however, it is more problematic. The defendants were not companies but individuals on a hunting trip. Mr. Tice, Mr. Simonson and Mr. Summers were walking through a field when a quail suddenly flew up. Tice and Simonson shot at it. Unfortunately, Summers had gotten out ahead of them, and one of the shots hit him in the eye. Like the *Sindell* plaintiffs, he could not meet the usual burden of proving which defendant caused the injury. As in *Sindell*, the court shifted the burden of proof. Both Tice and Simonson would have to compensate Summers unless either could prove that the shot was not his. And like the "innocent" DES manufacturers, the "innocent" hunter could not in fact exonerate himself by disproving causation. The type of shot fired by defendants Tice and Simonson was identical. The defendant whose shot missed, whoever he was, faced liability for harm that did not flow from his shooting (in an explanatory sense).

The problem with assigning liability this way is that unlike the DES manufacturers, the hunters probably did not believe they were engaged in a group enterprise governed by statistical probabilities. They would probably not divide up the game taken according to the percentage of shots fired, but would have tried to determine who had shot what bird, so that the outcome (pheasant under glass) would flow from individual action (a good shot). Assuming this was the arrangement, it is problematic to hold the hunters jointly liable for injuring Summers. And what if not two but 20 or 200 hunters happened to shoot in Summers' direction, still with the result that only one shot injured him? The larger the number of individuals involved, the more holding all of them responsible seems to be an unfair use of statistics. If individuals

did not act as a group, they should not be held liable as one.

On the other hand, each defendant did something predicted to injure the plaintiff, namely shoot in his direction at close range. It does not matter that in one case the predicted harm did not come to pass. In everyday usage, predictions are not guaranties, and knowing that a prediction might not come true does not destroy its usefulness. This logic is easily applied to legal liability. The (1) high likelihood that each hunter would predictably harm the plaintiff plus (2) the occurrence of the harm that was predicted was a sufficient causal connection, given (3) the absence of proof that one or the other hunter was *not* the cause. Thus, the court did make a properly individualized causal determination in the sense that it held individuals responsible for making the correct causal predictions. If Tice and Simonson should have known that they were likely to hit Summers, then they should not have shot. By pulling the trigger, they took a risk. It was then fair to hold them responsible if the predicted harm came to pass, especially since they had, at least technically, the right to provide another *explanation* if available (e.g., proof that the shot came from someone else's shotgun).

Can causation-as-prediction solve the Desert Traveler problem? In 1959, H. L. A. Hart and Antony Honoré proposed the NESS test, an approach to legal causation that incorporates causation-as-prediction.[14] It posits a set of events that, like a recipe, are sufficient to cause an outcome. It then asks if a candidate cause (such as a defendant's action) is a "Necessary Element" of this "Sufficient Set" (NESS). The NESS test explains the result in *Summers v. Tice* by looking at the set of all events that includes only one defendant's pulling the trigger. This set is enough (that is, sufficient) to injure Summers; there does not also need to have been a shot fired by the other defendant as long as all the other but-for causes (oxygen, no bulletproof vest, etc.) are in place.

The Desert Traveler problem presents a harder puzzle because the Traveler's enemies did not commit their bad acts simultaneously. Is the poisoning a NESS cause? No. The set of actions sufficient to poison the Traveler must include: "the Traveler drinks the poison." Since that did not happen, Poisoner is off the hook. Is Driller therefore the culprit? No. Hart and Honoré felt that saying a person killed someone is an implicit comparison with a situation in which the victim otherwise would have continued living, e.g., he had sufficient food, water, etc. Here, though, the Traveler could not have survived no matter what Driller did. Poisoner had already shortened his lifespan to virtually nothing by poisoning his only supply of water. Hart and Honoré concluded that the NESS test could not tell which enemy killed the Desert Traveler any more than the but-for test could. "Their mutual frustration of each other's plan precludes us from saying that either caused [the Traveler's] death."[15]

In 1980, J. L. Mackie tackled the Desert Traveler Problem by making a distinction between the fact "*That he* [the traveler] *died*, and *that he died of thirst.*" (Italics in original).[16] Death-by-thirst was the work of Driller. However, he did not cause the event of death. Death would have happened anyway, thanks to the poisoned water, which stood ready to kill the Traveler as what Mackie called the "reserve cause" of death.[17] Mackie explained: "In this way I would defend my judgment that the traveler's death—that is, the concrete event—was caused by the puncturing of the [water] can, but at the same time explain the doubts felt by lawyers (and others) about this case as due to the equally correct judgment that his dying—that is, the fact that he died on this journey—was not caused by the puncturing of the can."[18]

This distinction between a general outcome (dying) and the outcome-as-it-happened (dying of thirst) sharpens the question of what should trigger legal responsibility. Is it enough that the harm be generally predictable, as with these two near-certain predictions: "poisoning will kill the traveler" and "draining out the water will kill the traveler"? Or, must causation involve a specific explanation (the predicted poisoning did not happen; the predicted death by thirst did)?

It is logical to argue that only the second possibility (both prediction and explanation) actually identifies individual causal responsibility. This means that only Driller is legally responsible for killing the Traveler. The argument goes something like this: "Only Driller's actions actually *explain* what happened. Although Poisoner took steps to

kill the Traveler, the death-as-it-happened did not dirty his hands thanks to Driller's spilling out the poisoned water. People impose risks on others all the time, but incur no legal responsibility unless harm results from their actions, which is often simply a matter of bad luck. We cannot morally saddle people with responsibility for outcomes to which bad luck contributed without also excusing responsibility if, through good luck, no bad outcome follows from a wrong or foolish act. Although in the predictive sense, both enemies might have been equally blameworthy, it is more important that in the explanatory sense, they were not. Only causation-as-explanation can do what is particular to law, as distinguished from aggregate public policy determinations: namely, *hold people responsible for the things they did that made a difference in the world.* This rule should be followed even if in odd cases, such as the Desert Traveler, it may seem as if an equally bad actor gets off. In fact, both were not equally bad from the perspective of causation-as-explanation. What Driller did was worse because he actually brought about the death."

One problem with this argument, however, is that it produces the same result whether Driller spilled out good water or poisoned water. It fails to recognize what makes the Desert Traveler problem so interesting, namely the interlocking activities of Driller and Poisoner. Saying that the law must ignore Poisoner's involvement is intuitively unappealing.

Is there a way to hold both Driller and Poisoner responsible for Traveler's death? This certainly can be done if we see law as an instrument to increase general social welfare. Instead of using tort law to try to "correct" past injury by making the defendant pay for it, this view of the law looks instead to improve the future by imposing liability in a way that will lead to an increase in the total well-being of society. For instance, those who engage in high-risk activity (mining) should internalize the costs that the activity imposes on others (collapse of neighboring structures). Such a liability rule would have the goal of giving the risky actors incentives to take the proper level of care.[19] If the actors then took care (properly shored-up tunnels), there would be fewer injuries, making everyone better off. By this reasoning, if a rule holding people liable for outcomes that might be loosely connected with particular bad acts will discourage these acts, then the rule is good, even if some defendants might pay compensation for harms they did not cause. The Desert Traveler problem is a good candidate for this sort of forward-looking liability rule since there is no question that both poisoning and draining the water supply of a person about to enter the desert are activities that should be discouraged.

In short, on this view, a defendant does not need to be responsible for causing harm; it is enough if society would be better off with a decision holding the defendant liable. The problem with this solution, however, was discussed earlier. It too easily treats legal liability as a decision of policy or "practical politics" and lessens the ability of law to do justice in the individual case by demanding a close connection between the defendant's actions and the harmful outcome for which the defendant must pay.

So, we return to the initial question. Is there a way to hold both Poisoner and Driller liable for the death of the Desert Traveler without abandoning the requirement of causation? I suggest that predictive causation can do this. Both Driller and Poisoner did things that were *likely* to cause harm. The acts of drilling a hole in a water bottle and poisoning the water each impose a very high likelihood of causing death, and the Traveler did die. Even though there is a not a link between poisoning and the Traveler's death in the strongest explanatory but-for sense, there is such a link in the weaker sense of prediction.

Poisoner and Driller can be compared to Tice and Simonson. Each was found liable in *Summers v. Tice* because each had shot in the direction of Summers, an action predicted to cause injury. The court said that Tice and Simonson could each try to prove that he did not fire the shot that injured Summers' eye, but it was apparent that this proof was impossible. In the Desert Traveler, in contrast. Poisoner could exonerate himself by showing that Traveler died of thirst, which links the death to Driller. But why should this matter? The risky acts of Poisoner and Driller are as sufficient a causal link to the Traveler's injuries as the carelessness of Tice and Simonson was in *Summers v. Tice*

We therefore have two alternatives to the "policy/politics" welfare-enhancing approach when it comes to deciding whom to hold liable for killing

the Desert Traveler. There is a strict causal explanation that would hold only Driller liable, given that the death as it came about involved only dehydration, not poisoning. And there is causation-as-prediction, which would hold both Poisoner and Driller liable given that death was the clearly predictable result of what each of them did. Both these approaches to causation maintain a causal link between the defendant's act and the outcome that harms the plaintiff. In that way each reflects a commitment to individual justice over aggregate social welfare. These two tests, however, give different answers. Which should be used? This question brings us to the final sense of causation, the assignment of responsibility.

Legal Causation as Responsibility: The Substantial Factor Test

As noted, people in everyday life use causal thinking to decide whether to hold someone responsible for something, and they do this in the face of both factual uncertainties and disagreements over what responsibility means. A third approach to causation in the law applies this practice more directly to the question of legal responsibility. This is the "substantial factor" test, commonly used in U.S. courts.

This test equates causing something with being a "substantial factor" in making it happen. In the words of the official California jury instructions: "A substantial factor in causing harm is a factor that a reasonable person would consider to have contributed to the harm. It must be more than a remote or trivial factor. It does not have to be the only cause of the harm."[20] The second restatement of the law of torts explains: "The word 'substantial' is used to denote the fact that the defendant's conduct has such an effect in producing the harm as to lead reasonable men to regard it as a cause, *using that word in the popular sense in which there always lurks the idea of responsibility.*"[21] (italics added).

It is easy to object that this test lacks intellectual rigor. (Indeed, the commentators on the proposed third restatement of the law of torts suggest the "substantial factor" test be reserved for odd cases like the Desert Traveler.[22]) Just what is "substantial" is not well defined. And, is it right to take a factor that only "contributed" to an outcome and elevate it to the status of cause? Isn't the whole point to figure out what actually led to what? That is the "but-for" test's great virtue: it approaches certainty. But the but-for test, as we saw, demands too much certainty so that those, like Poisoner, whom justice and common sense would hold liable, are not found to be causes. The "substantial factor" test, in contrast, leaves quite a bit of leeway in identifying causes, since they only need to be non-remote, non-trivial contributing factors.

I believe that these "vices" of the substantial factor test turn out to be its virtues. First, because it is open-ended, it allows jurors to use either predictive or explanatory causation or a bit of both. Second, by specifying that a contributing factor is sufficient for liability, it avoids the logical tangles of something like the Desert Traveler problem where it is difficult to identify one cause. The actions of both enemies can be "substantial factors." Third, the test directs jurors to distinguish between substantial and remote or trivial factors. This similarly avoids the distraction of the many causes that can be identified by the but-for test. Fourth, the test's fuzziness allows the jury to focus on what is actually going on, namely the assignment of responsibility based on factual connections between act and outcome. As I discussed earlier, even if jurors might disagree on the theoretical question about whether responsibility arises as a matter of fact in the world or is merely imposed by society, they are almost surely familiar with the practical task of assigning responsibility—it is something people do all their lives. By calling on the jury to think about what a "reasonable person" would find causally contributed to harm, the substantial factor test invokes this social practice; it does not require resolution of the debate whether causation is discovered in the world or merely created by rules of law or policy.

A final argument in support of the substantial factor test comes from the field of moral philosophy. Some philosophers have observed that people do not—and should not—always apply moral reasons in the same ways. For instance, borrowing a book is a reason to return it to the lender. However, if the borrower learns that the book was stolen from the library, this reason disappears.

Indeed, now there is a reason not to return the book, but to give it to the library instead.[23] This view, known as moral particularism, does not deny that there *are* moral values (e.g., proper respect for others as persons, including respect for the property of others); it only denies that rules (e.g., "always return what you have borrowed") can tell us how these values should be vindicated in all situations. Rather, "one's main duty, in moral judgment, is to look really closely at the case before one."[24]

Moral particularism takes the argument in favor of the substantial factor test one step further. As noted, an advantage of the substantial factor test is that it does not require that on-going theoretical debates in the law abut the purpose and nature of causation be resolved. Moral particularism suggests that these debates not only need not be resolved but should not be resolved. This applies to the question of causation. Why is it important to understand causation correctly in any particular case? The reason is that causation tells us what impact the defendant's actions had on the world. Unless we know that, we cannot properly assign legal liability. Causation serves the *moral* purpose of making sure that law reflects actual responsibility for outcomes. Inasmuch as determining causation is a moral decision, moral particularism directs that it, too, be made by looking "really closely at the case before one."

So, who killed the Desert Traveler? In the predictive sense of causation, Poisoner and Driller both caused the death; in the strongest "but-for" sense of causation, neither did, and in the narrow sense of explaining the death-as-it-happened (from thirst), only Driller killed the Traveler. The substantial factor test allows the jury to choose among these possibilities, determining moral responsibility based on what actually happened. In practice, liability will never be decided simply on a stripped-down hypothetical such as the story that started this article. The function of a trial is to present the nuances and details of the situation. And the function of the jury is to look closely and carefully and assign responsibility for acts that caused harm. The substantial factor test directs jurors to bring all the tools and insights of ordinary experience to this vital task.

NOTES

1. J. A. McLaughlin, "Proximate Cause," 39 *Harvard Law Review* 149, 155, n. 25 (1925–26).
2. Antony Honoré, "Necessary and Sufficient Conditions," in *Philosophical Foundations of Tort Law* (David G. Owen, ed.) (Clarendon Press, 1997), 363, 375–376.
3. In the English case of *Bolton v. Stone* [1951] A.C. 850, the operator of a cricket ground was not held liable for injuries to a passerby in the neighboring street who was hit by a ball because the event was not reasonably predictable: "The test to be applied here is whether the risk of damage to a person on the road was so small that a reasonable man in the position of the [grounds operator], considering the matter from the point of view of safety, would have thought it right to refrain from taking steps to prevent the danger."
4. Two things must be true before a person can be identified as "responsible." First, she must be "eligible" for causal responsibility by meeting some standard of intellectual capacity that defines a responsible agent. Young children and some mentally incompetent people are not eligible. Second, even if a person is capable of responsibility, she had to be in a position to control her actions at the time of the relevant events. If Jack is pushed into a china display by a thug, Jack is not responsible for the broken dishes, and if Annette lights a match with no reason to know that the room was full of colorless, odorless natural gas, she is not responsible for causing the ensuing explosion. Note, in both cases, that the person's actions are part of the *explanation* of what happened: Jack did break the dishes and Annette did light the match. However, as things happened, neither is *responsible* for the outcome. Only after the thresholds of capacity and control are passed is the question of causation for ascribing responsibility presented.
5. The tort of trespass is a rare exception in which damages need not be proven.
6. *Black's Law Dictionary*, p. 212 (7th ed., 1999, West Group, St. Paul, Minn.).
7. In a similar situation, the highest court of New York did hold a hotel owner liable for the shooting of a visitor in its lobby, given the manager's knowledge of the history of crime in the building. *Nallan v. Helmsley-Spear, Inc*, 50 N.Y.2d 507 (1980).
8. 248 N.Y. 339 (1928).
9. See *Kingston v. Chicago & N.W. Ry.*, 191 Wis. 610 (1927), in which two fires destroyed the plaintiff's property, and the court held the party responsible for one of them liable for all the damages. The court noted: "To permit each of two wrongdoers to plead the wrong of the other as a defense to his own wrongdoing would permit both wrongdoers to escape and penalize the innocent party who has been damaged by their wrongful acts."

10. See, e.g., Moore, "The Metaphysics of Causal Intervention," 88 *California Law Review* 827 (2002).
11. I am not considering at this point a stronger position that hold that assigning legal responsibility is always a legal/policy decision. On this view, treating legal causation as if it were a fact simply hides the normative decision regarding where to assign responsibility. See, e.g., Jane Stapleton, "Perspectives on Causation," in *Oxford Essays in Jurisprudence,* 4th Series (J. Horder, ed.) (Oxford University Press, 2000).
12. 26 Cal.3d 588 (1980).
13. 33 Cal.2d 80 (1948).
14. H. L. A. Hart and A. M. Honoré, *Causation in the Law* (Oxford University Press, 1959); see A. M. Honoré, *Responsibility and Fault* (Oxford University Press, 1999), chapter 5.
15. *Causation in the Law* (2nd ed.) at 241.
16. J. L. Mackie, *The Cement of the Universe: A Study of Causation* (Oxford University Press, 1980), pp. 45–46.
17. *A Study of Causation* (Oxford University Press, 1980), p. 46. Honoré later came to agree with this analysis. Honoré, *Responsibility and Fault* (1999), 112–114.
18. J. L. Mackie, *The Cement of the Universe: A Study of Causation* (Oxford University Press, 1980), p. 46.
19. See Louis Kaplow and Steven Shavell, *Fairness versus Welfare* (Harvard University Press, 2002).
20. Jury Instruction No. 430, "Causation," Judicial Council of California Civil Jury Instructions, available at http://www.courtinfo.ca.gov/reference/documents/civiljuryinst.pdf, accessed September 2, 2004.
21. Restatement Second of Torts, section 431, comment a.
22. Restatement of the Law, Third, Torts: Liability for Physical Harm (Basic Principles) (Tentative Draft No. 2, March 25, 2002) section 26, comment j, Reporters' Notes.
23. Jonathan Dancy, *Moral Reasons* (Blackwell Publishers, 1993), 60–63.
24. *Id.* at 63.

REVIEW AND DISCUSSION QUESTIONS

1. What are the three everyday uses of causation, according to Shapiro?
2. In what sense is the but-for test a test that produces a causal "explanation"? In which situations is it prone to produce a result that seems to be at odds with intuitive concepts of justice?
3. How does making a prediction or giving a warning involve a determination of causation?
4. Shapiro claims it is better to use a test of causation-as-explanation than a test of causation-as-prediction when an individual (rather than a corporation) is alleged to have committed a tort. Why does Shapiro think this?
5. The article concludes that the "substantial factor" test is superior to other tests of causation because it allows jurors to use the ordinary practice of assigning responsibility when they are asked to assign legal responsibility. Does this argument hold for case in which a judge, rather than a jury, decides whether there is responsibility or not?
6. Who killed the desert traveler?

A Duty to Rescue?

Yania v. Bigan; Farwell v. Keaton; and *McFall v. Shimp*

The Bible holds up the Good Samaritan as a moral exemplar, but it is an ancient, if often criticized, principle of the common law that one has no affirmative legal duty to save the life of another person, if one is not responsible for his or her plight and if one has no prior legal responsibility for that person. In the following cases, the Supreme Court of Pennsylvania, the Supreme Court of Michigan, and a lower Pennsylvania court concern themselves with defendants who have been sued, in differing circumstances, for failing to save the life of another person. The conclusions they reach differ.

155 A. 2d 343 (1959); 240 N.W. 2d 217 (1976); No. 78-177711 (1978) 10th Penn. District, Allegheny County, respectively.

Yania v. Bigan

Justice Benjamin R. Jones: A bizarre and most unusual circumstance provides the background of this appeal.

On September 25, 1957 John E. Bigan was engaged in a coal strip-mining operation in Shade Township, Somerset County. On the property being stripped were large cuts or trenches created by Bigan when he removed the earthen overburden for the purpose of removing the coal underneath. One cut contained water 8 to 10 feet in depth with side walls or embankments 16 to 18 feet in height; at this cut Bigan had installed a pump to remove the water.

At approximately 4 P.M. on that date, Joseph F. Yania, the operator of another coal strip-mining operation, and one Boyd M. Ross went upon Bigan's property for the purpose of discussing a business matter with Bigan, and, while there, were asked by Bigan to aid him in starting the pump. Ross and Bigan entered the cut and stood at the point where the pump was located. Yania stood at the top of one of the cut's side walls and then jumped from the side wall—a height of 16 to 18 feet—into the water and was drowned.

Yania's widow, in her own right and on behalf of her three children, instituted wrongful death and survival actions against Bigan contending Bigan was responsible for Yania's death.

. . . Summarized, Bigan stands charged with threefold negligence: (1) by urging, enticing, taunting and inveigling Yania to jump into the water; (2) by failing to warn Yania of a dangerous condition on the land, i.e., the cut wherein lay 8 to 10 feet of water; (3) by failing to go to Yania's rescue after he had jumped into the water.[1]

. . . Appellant initially contends that Yania's descent from the high embankment into the water and the resulting death were caused "entirely" by the spoken words and blandishments of Bigan delivered at a distance from Yania. The complaint does not allege that Yania slipped or that he was pushed or that Bigan made any *physical* impact upon Yania. On the contrary, the only inference deducible from the facts alleged in the complaint is that Bigan, by the employment of cajolery and inveiglement, caused such a *mental* impact on Yania that the latter was deprived of his volition and freedom of choice and placed under a compulsion to jump into the water. Had Yania been a child of tender years or a person mentally deficient then it is conceivable that taunting and enticement could constitute actionable negligence if it resulted in harm. However, to contend that such conduct directed to an adult in full possession of all his mental faculties constitutes actionable negligence is not only without precedent but completely without merit.

. . . Appellant next urges that Bigan, as the possessor of the land, violated a duty owed to Yania in that his land contained a dangerous condition, i.e., the water-filled cut or trench, and he failed to warn Yania of such condition. Yania was a business invitee in that he entered upon the land for a common business purpose for the mutual benefit of Bigan and himself. . . .

The *only* condition on Bigan's land which could possibly have contributed in any manner to Yania's death was the water-filled cut with its high embankment. Of this condition there was neither concealment nor failure to warn, but, on the contrary, the complaint specifically avers that Bigan not only requested Yania and Boyd to assist him in starting the pump to remove the water from the cut but "led" them to the cut itself. If this cut possessed any potentiality of danger, such a condition was as obvious and apparent to Yania as to Bigan, both coal strip-mine operators. Under the circumstances herein depicted Bigan could not be held liable in this respect.

. . . Lastly, it is urged that Bigan failed to take the necessary steps to rescue Yania from the water. The mere fact that Bigan saw Yania in a position of peril in the water imposed upon him no legal, although a moral, obligation or duty to go to his rescue unless Bigan was legally responsible, in whole or in part, for placing Yania in the perilous position. . . . The language of this Court in *Brown v. French*, 104 Pa. 604, 607, 608, is apt: "If it appeared that the deceased, by his own carelessness, contributed in any degree to the accident which caused the loss of his life, the defendants ought not to have been held to

answer for the consequences resulting from that accident. . . . He voluntarily placed himself in the way of danger, and his death was the result of his own act. . . . That his undertaking was an exceedingly reckless and dangerous one, the event proves, but there was no one to blame for it but himself. He had the right to try the experiment, obviously dangerous as it was, but then also upon him rested the consequences of that experiment, and upon no one else; he may have been, and probably was, ignorant of the risk which he was taking upon himself, or knowing it, and trusting to his own skill, he may have regarded it as easily superable. But in either case, the result of his ignorance, or of his mistake, must rest with himself—and cannot be charged to the defendants." The complaint does not aver any facts which impose upon Bigan legal responsibility for placing Yania in the dangerous position in the water and, absent such legal responsibility, the law imposes on Bigan no duty of rescue.

Recognizing that the deceased Yania is entitled to the benefit of the presumption that he was exercising due care and extending to appellant the benefit of every well pleaded fact in this complaint and the fair inferences arising therefrom, yet we can reach but one conclusion: that Yania, a reasonable and prudent adult in full possession of all his mental faculties, undertook to perform an act which he knew or should have known was attended with more or less peril and it was the performance of that act and not any conduct upon Bigan's part which caused his unfortunate death.

Order affirmed.

Farwell v. Keaton

Justice Levin: On the evening of August 26, 1966, Siegrist and Farwell drove to a trailer lot to return an automobile which Siegrist had borrowed from a friend who worked there. While waiting for the friend to finish work, Siegrist and Farwell consumed some beer.

Two girls walked by the entrance to the lot. Siegrist and Farwell attempted to engage them in conversation; they left Farwell's car and followed the girls to a drive-in restaurant down the street.

The girls complained to their friends in the restaurant that they were being followed. Six boys chased Siegrist and Farwell back to the lot. Siegrist escaped unharmed, but Farwell was severely beaten. Siegrist found Farwell underneath his automobile in the lot. Ice was applied to Farwell's head. Siegrist then drove Farwell around for approximately two hours, stopping at a number of drive-in restaurants. Farwell went to sleep in the back seat of his car. Around midnight Siegrist drove the car to the home of Farwell's grandparents, parked it in the driveway, unsuccessfully attempted to rouse Farwell and left. Farwell's grandparents discovered him in the car the next morning and took him to the hospital. He died three days later of an epidural hematoma.

At trial, plaintiff contended that had Siegrist taken Farwell to the hospital, or had he notified someone of Farwell's condition and whereabouts, Farwell would not have died. A neurosurgeon testified that if a person in Farwell's condition is taken to a doctor before, or within half an hour after, consciousness is lost, there is an 85 to 88 percent chance of survival. Plaintiff testified that Siegrist told him that he knew Farwell was badly injured and that he should have done something.

The jury returned a verdict for plaintiff and awarded $15,000 in damages. The Court of Appeals reversed, finding that Siegrist had not assumed the duty of obtaining aid for Farwell and that he neither knew nor should have known of the need for medical treatment. . . .

Siegrist contends that he is not liable for failure to obtain medical assistance for Farwell because he had no duty to do so.

Courts have been slow to recognize a duty to render aid to a person in peril. Where such a duty has been found, it has been predicated upon the existence of a special relationship between the parties; in such a case, if defendant knew or should have known of the other person's peril, he is required to render reasonable care under all the circumstances. . . .

Farwell and Siegrist were companions on a social venture. Implicit in such a common undertaking is the understanding that one will render assistance to the other when he is in peril if he can do so without endangering himself. Siegrist knew or should have known when he left Farwell, who was badly beaten and unconscious, in the back seat of his car that no one would find him before morning. Under these circumstances, to say that Siegrist had no duty to obtain medical assistance or at least to notify someone of Farwell's condition and whereabouts would be "shocking to humanitarian considerations" and fly in the face of "the commonly accepted code of social conduct." "[C]ourts will find a duty where, in general, reasonable men would recognize it and agree that it exists."

Farwell and Siegrist were companions engaged in a common undertaking; there was a special relationship between the parties. Because Siegrist knew or should have known of the peril Farwell was in and could render assistance without endangering himself he had an affirmative duty to come to Farwell's aid.

Justice Fitzgerald (dissenting): The unfortunate death of Richard Farwell prompted this wrongful death action brought by his father against the defendant, David Siegrist, a friend who had accompanied Farwell during the evening in which the decedent received injuries which ultimately caused his death three days later. The question before us is whether the defendant, considering his relationship with the decedent and the activity they jointly experienced on the evening of August 26–27, 1966, by his conduct voluntarily or otherwise assumed, or should have assumed, the duty of rendering medical or other assistance to the deceased. We find that defendant had no obligation to assume, nor did he assume, such a duty. . . .

Defendant did not voluntarily assume the duty of caring for the decedent's safety. Nor did the circumstances which existed on the evening of August 26, 1966, impose such a duty. Testimony revealed that only a qualified physician would have reason to suspect that Farwell had suffered an injury which required immediate medical attention. The decedent never complained of pain and, in fact, had expressed a desire to retaliate against his attackers. Defendant's inability to arouse the decedent upon arriving at his grandparents' home does not permit us to infer, as does plaintiff, that defendant knew or should have known that the deceased was seriously injured. While it might have been more prudent for the defendant to insure that the decedent was safely in the house prior to leaving, we cannot say that defendant acted unreasonably in permitting Farwell to spend the night asleep in the back seat of his car.

The close relationship between defendant and the decedent is said to establish a legal duty upon defendant to obtain assistance for the decedent. No authority is cited for this proposition other than the public policy observation that the interest of society would be benefited if its members were required to assist one another. This is not the appropriate case to establish a standard of conduct requiring one to legally assume the duty of insuring the safety of another. Recognizing that legal commentaries have expressed moral outrage at those decisions which permit one to refuse aid to another whose life may be in peril, we cannot say that, considering the relationship between these two parties and the existing circumstances, defendant acted in an unreasonable manner.

Plaintiff believes that a legal duty to aid others should exist where such assistance greatly benefits society and only a reasonable burden is imposed upon those in a position to help. He contends further that the determination of the existence of a duty must rest with the jury where questions of foreseeability and the relationship of the parties are primary considerations.

It is clear that defendant's nonfeasance, or the "passive inaction or a failure to take steps to protect [the decedent] from harm" is urged as being the proximate cause of Farwell's death. We must reject plaintiff's proposition which elevates a moral obligation to the level of a legal duty where, as here, the facts within defendant's knowledge in no way indicated that immediate medical attention was necessary and the relationship between the parties imposes no affirmative duty to render assistance. . . . The posture of this case does not permit us to create a legal duty upon one to render assistance to another injured or imperiled party where the initial injury was not caused by the person upon whom the duty is sought to be imposed.

McFall v. Shimp

Judge Flaherty: The Plaintiff, Robert McFall, suffers from a rare bone marrow disease and the prognosis for his survival is very dim, unless he receives a bone marrow transplant from a compatible donor. Finding a compatible donor is a very difficult task, and limited to a selection among close relatives. After a search and certain tests, it has been determined that only the Defendant is suitable as a donor. The defendant refuses to submit to the necessary transplant, and before the Court is a request for a preliminary injunction which seeks to compel the defendant to submit to further tests, and, eventually, the bone marrow transplant.

Although a diligent search has produced no authority, the Plaintiff cites the ancient statute of King Edward I, St. Westminster 2, 13 Ed. I, c 24, pointing out, as is the case, that this Court is a successor to the English courts of Chancery and derives power from this statute, almost 700 years old. The question posed by the Plaintiff is that, in order to save the life of one of its members by the only means available, may society infringe upon one's absolute right to his "bodily security"?

The common law has consistently held to a rule which provides that one human being is under no legal compulsion to give aid or to take action to save that human being or to rescue. A great deal has been written regarding this rule which, on the surface, appears to be revolting in a moral sense. Introspection, however, will demonstrate that the rule is founded upon the very essence of our free society. It is noteworthy that counsel for the Plaintiff has cited authority which has developed in other societies in support of the Plaintiff's request in this instance. Our society, contrary to many others, has as its first principle, the respect for the individual, and that society and government exist to protect the individual from being invaded and hurt by another. Many societies adopt a contrary view which has the individual existing to serve the society as a whole. In preserving such a society as we have it is bound to happen that great moral conflicts will arise and will appear harsh in a given instance. In this case, the chancellor is being asked to force one member of society to undergo a medical procedure which would provide that part of that individual's body would be removed from him and given to another so that the other could live. Morally, this decision rests with the Defendant, and, in the view of the Court, the refusal of the Defendant is morally indefensible. For our law to *compel* the Defendant to submit to an intrusion of his body would change the very concept and principle upon which our society is founded. To do so would defeat the sanctity of the individual, and would impose a rule which would know no limits, and one could not imagine where the line would be drawn. This request is not to be compared with an action at law for damages, but rather is an action in equity before a Chancellor, which, in the ultimate, if granted, would require the [forcible] submission to the medical procedure. For a society, which respects the rights of *one* individual, to sink its teeth into the jugular vein or neck of one of its members and suck from it sustenance for *another* member, is revolting to our hardwrought concepts of jurisprudence. [Forcible] extraction of living body tissue causes revulsion to the judicial mind. Such would raise the spectre of the swastika and the inquisition, reminiscent of the horrors this portends.

This court makes no comment on the law regarding the Plaintiff's right in an action at law for damages, but has no alternative but to deny the requested equitable relief. An Order will be entered denying the request for a preliminary injunction.

NOTE

1. So far as the record is concerned we must treat the 33-year-old Yania as in full possession of his mental faculties at the time he jumped.

REVIEW AND DISCUSSION QUESTIONS

1. Did Bigan cause Yania's death? Was he morally responsible for it? Should he be held legally liable for it?
2. According to the court, why did Bigan have no legal duty to rescue Yania? In your view, should he have such a duty?
3. On what grounds does the court hold David Siegrist responsible for not saving Richard Farwell's life? Is it possible to make Justice Levin's decision consistent with the verdict in *Yania v. Bigan?*
4. On what grounds does Justice Fitzgerald dissent in the *Farwell* case? What decision would you have reached?
5. Why does Judge Flaherty refuse to compel Shimp to submit to a bone marrow transplant? Does Shimp have a moral duty to do so? Legality aside, would compelling Shimp to do so be morally right? Does it make a difference that McFall and Shimp are related?
6. Is the *McFall* decision consistent with the *Farwell* case? Suppose a different court held Bigan liable for Yania's death but did not compel Shimp to undergo the operation. Would this be consistent?
7. How ought the law to handle these three cases? Should people be legally obligated to be Good Samaritans?

Wrongful Life and Wrongful Birth

Berman v. Allan

A basic element of negligence or any other tort is that the plaintiff suffered some wrong, harm, or injury. But can life itself be a harm? That question is addressed in this case, in which suit was brought, in part, on behalf of an infant suffering from Down's syndrome against her mother's doctor. The doctor was not responsible for the infant's affliction, but he was negligent in failing to tell the mother, Shirley Berman, about amniocentesis, a procedure that would have indicated this birth defect early enough in Mrs. Berman's pregnancy to have enabled her to have terminated it. On appeal, the Supreme Court of New Jersey rejected the idea that "wrongful life" can be a valid cause of legal action. But it also broke with precedent and held that the parents could claim damages in their own right for "wrongful birth." In his partially dissenting opinion, Justice Handler argues that the child herself has suffered an injury that calls for compensation.

The opinion of the court was delivered by PASHMAN, J.:

1

In *Gleitman v. Cosgrove*, decided 12 years ago, this Court refused to recognize as valid causes of action either a claim for "wrongful life" asserted on behalf of a physically deformed infant or a claim for "wrongful birth" put forth by the infant's parents. Both prayers for relief were premised upon the allegation that had the physician treating Mrs. Gleitman during her pregnancy followed standard medical practice, an abortion would have been procured and the child would never have come into existence. In this case, we are called upon to assess the continued validity of both of our holdings in *Gleitman.*

On September 11, 1975, Paul and Shirley Berman, suing both in their own names and as Guardians *ad litem* for their infant daughter Sharon, instituted the present malpractice action against Ronald Allan and Michael Attardi, medical doctors licensed by the State of New Jersey.

. . . At the time of her pregnancy, Mrs. Berman was 38 years of age. On November 3, Sharon was

From 404 A.2d 8 (1979). Some citations omitted.

born afflicted with Down's Syndrome—a genetic defect commonly referred to as mongolism.

Plaintiffs allege that defendants deviated from accepted medical standards by failing to inform Mrs. Berman during her pregnancy of the existence of a procedure known as amniocentesis. This procedure involves the insertion of a long needle into a mother's uterus and the removal therefrom of a sample of amniotic fluid containing living fetal cells. Through "karyotype analysis"—a procedure in which the number and structure of the cells' chromosomes are examined—the sex of the fetus as well as the presence of gross chromosomal defects can be detected. . . .

Due to Mrs. Berman's age at the time of her conception, plaintiffs contend that the risk that her child, if born, would be afflicted with Down's Syndrome was sufficiently great that sound medical practice at the time of pregnancy required defendants to inform her both of this risk and the availability of amniocentesis as a method of determining whether in her particular case that risk would come to fruition. Had defendants so informed Mrs. Berman, the complaint continues, she would have submitted to the amniocentesis procedure, discovered that the child, if born, would suffer from Down's Syndrome, and had the fetus aborted.

As a result of defendants' alleged negligence, the infant Sharon, through her Guardian *ad litem*, seeks compensation for the physical and emotional pain and suffering which she will endure throughout life because of her mongoloid condition. Mr. and Mrs. Berman, the child's parents, request damages in their own right both for the emotional anguish which they have experienced and will continue to experience on account of Sharon's birth defect, and the medical and other costs which they will incur in order to properly raise, educate and supervise the child.

On November 4, 1977, the trial judge granted summary judgment in favor of defendants on the ground that plaintiffs had failed to state any actionable claim for relief. . . .

2

The claim for damages asserted on behalf of the infant Sharon has aptly been labeled a cause of action grounded upon "wrongful life." Sharon does not contend that absent defendants' negligence she would have come into the world in a normal and healthy state. There is no suggestion in either the pleadings below or the medical literature which we have scrutinized that any therapy could have been prescribed which would have decreased the risk that, upon birth, Sharon would suffer from Down's Syndrome. Rather, the gist of the infant's complaint is that had defendants informed her mother of the availability of amniocentesis, Sharon would never have come into existence.

As such, this case presents issues different from those involved in malpractice actions where a plaintiff asserts that a defendant's deviation from sound medical practices *increased* the probability that an infant would be born with defects. . . . Nor are we here confronted with a situation in which an individual's negligence while a child was in gestation caused what otherwise would have been a normal and healthy child to come into the world in an impaired condition. . . . Here, defendants' alleged negligence neither caused the mongoloid condition nor increased the risk that such a condition would occur. In the words of the *Gleitman* majority, "the infant plaintiff [asserts]. . . . not that [she] should have been born without defects but [rather] that [she] should not have been born at all." . . . In essence, Sharon claims that her very life is "wrongful."

The *Gleitman* majority refused to recognize as valid a cause of action predicated upon wrongful life. Its main reason for so holding was that damages would be impossible to ascertain. . . .

The primary purpose of tort law is that of compensating plaintiffs for the injuries they have suffered wrongfully at the hands of others. As such, damages are ordinarily computed by "comparing the condition plaintiff would have been in, had the defendants not been negligent, with plaintiff's impaired condition as a result of the negligence." . . . In the case of a claim predicated upon wrongful life, such a computation would require the trier of fact to measure the difference in value between life in an impaired condition and the "utter void of nonexistence." . . . Such an endeavor, however, is literally impossible. As Chief Justice Weintraub noted, man, "who knows nothing of death or nothingness," simply cannot affix a price tag to non-life. . . .

Nevertheless, although relevant to our determination, we would be extremely reluctant today to deny the validity of Sharon's complaint solely because damages are difficult to ascertain. The courts of this and other jurisdictions have long held that where a wrong itself is of such a nature as to preclude the computation of damages with precise exactitude, it would be a "perversion of fundamental principles of justice to deny all relief to the injured [party], and thereby relieve the wrongdoer from making any amend for his acts." . . . To be sure, damages may not be determined by mere speculation or guess and, as defendants emphasize, placing a value upon non-life is not simply difficult—it is humanly impossible. Nonetheless, were the *measure* of damages our sole concern, it is possible that some judicial remedy could be fashioned which would redress plaintiff, if only in part, for injuries suffered. . . .

Difficulty in the *measure* of damages is not, however, our sole or even primary concern. Although we conclude, as did the *Gleitman* majority, that Sharon has failed to state an actionable claim for relief, we base our result upon a different premise—that Sharon has not suffered any damage cognizable at law by being brought into existence. . . .

One of the most deeply held beliefs of our society is that life—whether experienced with or without a major physical handicap—is more precious than non-life. . . . Concrete manifestations of this belief are not difficult to discover. The documents which set forth the principles upon which our society is founded are replete with references to the sanctity of life. The federal constitution characterizes life as one of three fundamental rights of which no man can be deprived without due process of law. *U.S. Const.*, Amends V and XIV. Our own state constitution proclaims that the "enjoying and defending [of] life" is a natural right. . . . The Declaration of Independence states that the primacy of man's "unalienable" right to life is a "self-evident truth." Nowhere in these documents is there to be found an indication that the lives of persons suffering from physical handicaps are to be less cherished than those of non-handicapped human beings.

State legislatures—and thus the people as a whole—have universally reserved the most severe criminal penalties for individuals who have unjustifiably deprived others of life. Indeed, so valued is this commodity that even one who has committed first degree murder cannot be sentenced to death unless he is accorded special procedural protections in addition to those given all criminal defendants. . . . Moreover, it appears that execution is constitutionally impermissible unless the crime which a defendant has perpetrated was one which involved the taking of another's life. . . . Again, these procedural protections and penalties do not vary according to the presence or absence of physical deformities in the victim or defendant. It is life itself that is jealously safeguarded, not life in a perfect state.

Finally, we would be remiss if we did not take judicial notice of the high esteem which our society accords to those involved in the medical profession. The reason for this is clear. Physicians are the preservers of life.

No man is perfect. Each of us suffers from some ailments or defects, whether major or minor, which make impossible participation in all the activities the world has to offer. But our lives are not thereby rendered less precious than those of others whose defects are less pervasive or less severe.

We recognize that as a mongoloid child, Sharon's abilities will be more circumscribed than those of normal, healthy children and that she, unlike them, will experience a great deal of physical and emotional pain and anguish. We sympathize with her plight. We cannot, however, say that she would have been better off had she never been brought into the world. Notwithstanding her affliction with Down's Syndrome, Sharon, by virtue of her birth, will be able to love and be loved and to experience happiness and pleasure—emotions which are truly the essence of life and which are far more valuable than the suffering she may endure. To rule otherwise would require us to disavow the basic assumption upon which our society is based. This we cannot do.

Accordingly, we hold that Sharon has failed to state a valid cause of action founded upon "wrongful life."

3

The validity of the parents' claim for relief calls into play considerations different from those involved in the infant's complaint. As in the case of the infant, Mr. and Mrs. Berman do not assert that

defendants increased the risk that Sharon, if born, would be afflicted with Down's Syndrome. Rather, at bottom, they allege that they were tortiously injured because Mrs. Berman was deprived of the option of making a meaningful decision as to whether to abort the fetus. . . . They thus claim that Sharon's "birth"—as opposed to her "life"—was wrongful.

Two items of damage are requested in order to redress this allegedly tortious injury: (1) the medical and other costs that will be incurred in order to properly raise, supervise and educate the child; and (2) compensation for the emotional anguish that has been and will continue to be experienced on account of Sharon's condition.

The *Gleitman* majority refused to recognize as valid a cause of action grounded upon wrongful birth. Two reasons underlay its determination. The first related to measure of damages should such a claim be allowed. In this view,

> In order to determine [the parents'] compensatory damages a court would have to evaluate the denial to them of the intangible, unmeasurable, and complex human benefits of motherhood and fatherhood and weigh these against the alleged emotional and money injuries. Such a proposed weighing is . . . impossible to perform. . . .

Second, even though the Court's opinion was premised upon the assumption that Mrs. Gleitman could have legally secured an abortion, the majority concluded that "substantial [public] policy reasons" precluded the judicial allowance of tort damages "for the denial of the opportunity to take an embryonic life." . . .

In light of changes in the law which have occurred in the 12 years since *Gleitman* was decided, the second ground relied upon by the *Gleitman* majority can no longer stand in the way of judicial recognition of a cause of action founded upon wrongful birth. The Supreme Court's ruling in *Roe v. Wade* . . . clearly establishes that a woman possesses a constitutional right to decide whether her fetus should be aborted, at least during the first trimester of pregnancy. Public policy now supports, rather than militates against, the proposition that she not be impermissibly denied a meaningful opportunity to make that decision.

As in all other cases of tortious injury, a physician whose negligence has deprived a mother of this opportunity should be required to make amends for the damage which he has proximately caused. Any other ruling would in effect immunize from liability those in the medical field providing inadequate guidance to persons who would choose to exercise their constitutional right to abort fetuses which, if born, would suffer from genetic defects. . . . Accordingly, we hold that a cause of action founded upon wrongful birth is a legally cognizable claim.

Troublesome, however, is the measure of damages. As noted earlier, the first item sought to be recompensed is the medical and other expenses that will be incurred in order to properly raise, educate and supervise the child. Although these costs were "caused" by defendants' negligence in the sense that but for the failure to inform, the child would not have come into existence, we conclude that this item of damage should not be recoverable. In essence, Mr. and Mrs. Berman desire to retain all the benefits inhering in the birth of the child—i.e., the love and joy they will experience as parents—while saddling defendants with the enormous expenses attendant upon her rearing. Under the facts and circumstances here alleged, we find that such an award would be wholly disproportionate to the culpability involved, and that allowance of such a recovery would both constitute a windfall to the parents and place too unreasonable a financial burden upon physicians. . . .

The parents' claim for emotional damages stands upon a different footing. In failing to inform Mrs. Berman of the availability of amniocentesis, defendants directly deprived her—and, derivatively, her husband—of the option to accept or reject a parental relationship with the child and thus caused them to experience mental and emotional anguish upon their realization that they had given birth to a child afflicted with Down's Syndrome. . . .

Unlike the *Gleitman* majority, we do not feel that placing a monetary value upon the emotional suffering that Mr. and Mrs. Berman have and will continue to experience is an impossible task for the trier of fact. In the 12 years that have elapsed since *Gleitman* was decided, courts have come to recognize that mental and emotional distress is just as

"real" as physical pain, and that its valuation is no more difficult. Consequently, damages for such distress have been ruled allowable in an increasing number of contexts. . . . Moreover, . . . to deny Mr. and Mrs. Berman redress for their injuries merely because damages cannot be measured with precise exactitude would constitute a perversion of fundamental principles of justice. . . .

Consequently, we hold that Mr. and Mrs. Berman have stated actionable claims for relief. Should their allegations be proven at trial, they are entitled to be recompensed for the mental and emotional anguish they have suffered and will continue to suffer on account of Sharon's condition.

Accordingly, the judgment of the trial court is affirmed in part and reversed in part, and this case remanded for a plenary trial.

HANDLER, J., concurring in part and dissenting in part:

1

. . . The Court now recognizes that the parents of the impaired child have a cause of action for the doctors' breach of duty to render competent medical advice and services and that they are entitled to compensation for their mental and emotional suffering over the birth of their damaged child. I agree with this. However, I hold to a somewhat broader view of mental and emotional injury in these circumstances and would also include as an element of these damages impaired parenthood or parental capacity.

The Court does not, in its opinion, recognize as sustainable a cause of action on behalf of the child. On this, I differ. The child, in my view, was owed directly, during its gestation, a duty of reasonable care from the same physicians who undertook to care for its mother—then expectant—and that duty, to render complete and competent medical advice, was seriously breached. The child, concededly, did not become defective because of the physicians' dereliction; nevertheless, it suffered a form of injury or loss in having been born of parents whose parental capacity may have been substantially diminished by the negligence of their doctors. This is a loss to the child which should be recompensed. For these reasons I concur in part and dissent in part from the opinion of the Court.

2

It is important to have a clear picture of the claims which are asserted on behalf of the respective plaintiffs. . . .

Without doubt, expectant parents, kept in ignorance of severe and permanent defects affecting their unborn child, suffer greatly when the awful truth dawns upon them with the birth of the child. Human experience has told each of us, personally or vicariously, something of this anguish. Parents of such a child experience a welter of negative feelings—bewilderment, guilt, remorse and anguish—as well as anger, depression and despair. . . . When such a tragedy comes without warning these terrible emotions are bound to be felt even more deeply. "Novelty shock" may well exacerbate the suffering. . . . This, I believe, is the crux of the wrong done in this case. Through the failure of the doctors to advise an expectant mother, and father, of the likelihood or certainty of the birth of a mongoloid child, the parents were given no opportunity to cushion the blow, mute the hurt, or prepare themselves as parents for the birth of their seriously impaired child. Their injury is real and palpable. . . .

Because of the unique nature of the tort, involving as it does the denial of the opportunity to decide whether to become the parents of a handicapped child, the suffering of the parents assumes another, important dimension. There should be recognized in the stressful setting of this case the reality of moral injury. Such injury may be thought of as the deprivation of moral initiative and ethical choice. . . . Persons, confronted with the awesome decision of whether or not to allow the birth of a defective child, face a moral dilemma of enormous consequence. They deal with a profound moral problem. . . . To be denied the opportunity—indeed, the right—to apply one's own moral values in reaching that decision is a serious, irreversible wrong. . . . Shorn of ethical choice in bringing into the world a defective human being, some individuals will be torn by moral conflict. Moral suffering in this sense may be felt keenly by a person

who, as a matter of personal conscience, would choose not to allow the birth of such a child.... The moral affront, however, is not diminished because the parents, if given the choice, would have permitted the birth of the child. The crucial moral decision, which was theirs to make, was denied them.

A full perception of the mental, emotional—and, I add, moral—suffering of parents in this situation reveals another aspect of their loss. Mental, emotional and moral suffering can involve diminished parental capacity. Such incapacity of the mother and father *qua* parents is brought about by the wrongful denial of a reasonable opportunity to learn of and anticipate the birth of a child with permanent defects, and to prepare for the heavy obligations entailed in rearing so unfortunate an individual. Such parents may experience great difficulty adjusting to their fate and accepting the child's impairment as nature's verdict.... Individuals suffering this form of parental incapacity or dysfunction are denied to a great extent the fuller joys, satisfaction and pride which comes with successful and effective parenting. This may endure for some time during the early developmental years of the child.... Impaired parenthood, so understood, constitutes another dimension of the injury and loss suffered by plaintiffs in this case. In this sense, impaired parenthood, together with mental and emotional and moral suffering, should be recognized and compensated as elements of damages.

3

The Court in this case, as in *Gleitman* before it, fails to accord a cause of action to the afflicted infant plaintiff. This denial, I most respectfully urge, is wrong....

An adequate comprehension of the infant's claims under these circumstances starts with the realization that the infant has come into this world and is here, encumbered by an injury attributable to the malpractice of the doctors. That injury does not consist of the child's afflicted condition; her affliction was not the doctor's doing. Rather, the injury consists of a diminished childhood in being born of parents kept ignorant of her defective state while unborn and who, in that account, were less fit to accept and assume their parental responsibilities. The frightful weight of the child's natural handicap has been made more burdensome by defendants' negligence because her parents' capacity has been impaired; they are less able to cope with the extra-heavy parental obligations uniquely involved in providing a child so afflicted with the unfaltering love, constant devotion and extraordinary care such a child specially requires.

There has been some judicial appreciation of the notion that a diminished childhood may be the consequence of impaired parental capacity. In *Berger v. Weber*... the Court recognized a cause of action in a suit brought by a mentally retarded child to recover damages for the loss of society, companionship, love and affection of her mother as a result of injuries sustained by the mother in an auto accident. The mother had sustained both physical and psychological injuries and could no longer continue to administer to the peculiar needs of her retarded daughter. The Court upheld a cause of action for such losses when a parent is "severely" injured. It recognized that the loss of parental guidance and training can have a severe impact on a child's development and personality and found the reasons for denying a cause of action, such as lack of precedent, uncertainty of damages, possibility of double recovery and potential for a multiplicity of suits, unpersuasive. The Court also thought it anomalous to say that a child suffers a compensable loss in a wrongful death action when the parent is killed but not when a parent is injured so severely that he or she cannot perform the parental function. It concluded that the magnitude of the child's loss outweighs the factors which would militate against allowing recovery....

Plausibly, the child's injury and loss in the form of diminished childhood can be viewed as a derivative claim based solely on the parents' injury. The great majority of jurisdictions which have considered a child's cause of action for damages as derivative from injuries negligently inflicted on a parent have generally rejected such claims. This, in fact, was the unanimous holding of the Court in *Russell v. Salem Transportation Co., Inc.*.... The reasons for denying recognition of such a cause of action in *Russell* were several. The Court felt that the damages to the child were remote and speculative and feared a substantial accretion of liability against a tortfeasor from a single transaction, a possibility

of a double recovery in that consequential damages to children from a parent's accident are frequently already compensated in jury awards and conflicts within families in relation to the apportionment of settlement offers. . . . According to W. Prosser, *Law of Torts* (4th ed. 1971) § 125, pp. 898–899 (fns. omit.):

> It is not easy to understand and appreciate this reluctance to compensate the child who has been deprived of the care, companionship and education of his mother, or for that matter his father, through the defendant's negligence. This is surely a genuine injury, and a serious one, which has received a great deal more sympathy from the legal writers than from the judges. There is of course the . . . problem of preventing double compensation . . . , since the child will to some extent benefit by any sum recovered by the injured parent; but it is quite evident that this will not and cannot recompense him for all that he has lost. The obstacles in the way of satisfactory limitation of recovery are no greater than in the case of the wife [suing for loss of consortium]. As has been said even by one court which considered itself forced to deny recovery, it is difficult "on the basis of natural justice to reach the conclusion that this type of action will not lie." It is particularly difficult when recovery is permitted to the wife, but denied to the child.

Solution to this legal conundrum—whether and how to protect the infant's interests—should not turn on labels or definitions. It is to be emphasized in this case that the doctors' medical malpractice encompasses the child. . . . Their negligence consists of the failure to render proper advice to Mrs. Berman as an *expectant* mother. Indisputably in this relationship the doctors were caring for the unborn child as well as the mother; the duty they owed to Mrs. Berman enveloped a duty to the unborn child. The breach of that duty affects both. "The risk created by a negligent act of one who stands in a physician-patient relationship is of enormous consequence to mother and child."

REVIEW AND DISCUSSION QUESTIONS

1. Why does the court reject the "wrongful life" portion of the suit? Do you agree with Justice Pashman's argument?
2. Can someone's life and future prospects be so terrible that he or she would be better off not having been born? Do you think this was true in Sharon Berman's case?
3. The earlier *Gleitman* case refused to recognize "wrongful birth" as a cause of action for two reasons. How does the court now respond to those two considerations? Do you agree that parents can be wronged by not being given medical information that would have led them to obtain an abortion?
4. What does Justice Handler mean by "impaired parenthood or parental capacity"? Were the Bermans damaged in this way? Should such impairment, if it exists, be the basis for compensation?
5. According to Justice Handler, what exactly is the injury that young Sharon Berman has suffered? How persuasive do you find his reasoning?

CHAPTER 17

Private Ownership: The Law of Property

Property law is a large part of the law simply because private property is such a fundamental feature of our socioeconomic system. Inevitably, philosophical disagreements over the nature and justification of private property constitute the backdrop to legal debates over the adjudication of particular cases and the interpretation of property law in general. One philosophical tradition maintains that private property is a natural right and that the only legitimate role of positive property law is to uphold and safeguard that fundamental right. This perspective is represented by the selection from John Locke's *Second Treatise of Government.* A rival philosophical tradition rejects the notion of an unbridled natural right to property in favor of a more social perspective on property and a more utilitarian attitude toward it. This tradition emphasizes that property is a socioeconomic institution, with different societies having different systems of property, that is, different rules for determining what constitutes property and for governing its acquisition and transfer. In this view, the courts should interpret and shape property law so as to benefit society. This philosophical perspective is represented here by the selection from Morris Cohen. Although he accepts the idea of natural rights, Cohen maintains that the state can be justified in regulating and even in confiscating private property.

With the rise of the information economy, the importance of intellectual property has grown as have legal disputes over intellectual property rights. Richard Posner explores the nature of intellectual property and the tensions that lie at the heart of the legal rules governing it. The cases that follow raise basic questions about private property. *Haslem v. Lockwood* deals with the initial acquisition of private property, that is, the circumstances in which a previously unowned item becomes property. The *Pennsylvania Coal* case addresses the constitutional legitimacy of government-imposed restrictions on the use of private property, whereas *Kelo* concerns the related question of eminent domain—two issues that divide the two schools of thought distinguished earlier.

Property

John Locke

Few, if any, political theorists have had as much influence on the thinking of the framers of the American legal and constitutional order as John Locke did. Locke believed that legitimate government derives its powers from the consent of the governed and that such consent is best understood by first imagining people in a pre-political "state of nature." There, Locke thought, people would accept certain restrictions on their natural rights in order to gain the protection and other benefits of government.

In the following excerpt from his *Second Treatise of Government*, Locke argues that, even though God gave the world to humanity in common, in the state of nature people have a natural right to property based on their labor. This right to acquire property, though, is subject to certain limitations: "enough and as good" must be left for others, and we have no title to what we do not use or allow to spoil. The introduction of money, however, enables us to overcome these limitations and permits people legitimately to heap up wealth beyond what they can immediately use. In this way, Locke squares inequalities of wealth and possession with the initial equality of the state of nature.

OF THE STATE OF NATURE

To understand political power aright, and derive it from its original, we must consider what state all men are naturally in, and that is a state of perfect freedom to order their actions and dispose of their possessions and persons as they think fit, within the bounds of the law of nature, without asking leave, or depending upon the will of any other man.

A state also of equality, wherein all the power and jurisdiction is reciprocal, no one having more than another; there being nothing more evident than that creatures of the same species and rank, promiscuously born to all the same advantages of nature, and the use of the same faculties, should also be equal one amongst another without subordination or subjection, unless the Lord and Master of them all should by any manifest declaration of His will set one above another, and confer on him by an evident and clear appointment an undoubted right to domination and sovereignty.

But though this be a state of liberty, yet it is not a state of license; though man in that state have an uncontrollable liberty to dispose of his person or possessions, yet he has not liberty to destroy himself, or so much as any creature in his possession, but where some nobler use than its bare preservation calls for it. The state of nature has a law of nature to govern it, which obliges everyone; and reason, which is that law, teaches all mankind who will but consult it, that, being all equal and independent, no one ought to harm another in his life, health, liberty, or possessions. For men being all the workmanship of one omnipotent and infinitely wise Maker—all the servants of one sovereign Master, sent into the world by His order, and about His business—they are His property, whose workmanship they are, made to last during His, not one another's pleasure; and being furnished with like faculties, sharing all in one community of nature, there cannot be supposed any such subordination among us, that may authorize us to destroy one another, as if we were made for one another's uses, as the inferior ranks of creatures are for ours. Everyone, as he is bound to preserve himself, and not to quit his station willfully, so, by the like reason, when his own preservation comes not in competition, ought he, as much as he can, to preserve the rest of mankind, and not, unless it be to do justice on an offender, take away or impair the life, or what tends to the preservation of the life, the liberty, health, limb, or goods of another.

And that all men may be restrained from invading others' rights, and from doing hurt to one another, and the law of nature be observed, which willeth the peace and preservation of all mankind,

the execution of the law of nature is in that state put into every man's hand, whereby everyone has a right to punish the transgressors of that law to such a degree as may hinder its violation. For the law of nature would, as all other laws that concern men in this world, be in vain if there were nobody that, in the state of nature, had a power to execute that law, and thereby preserve the innocent and restrain offenders. And if anyone in the state of nature may punish another for any evil he has done, everyone may do so. For in that state of perfect equality, where naturally there is no superiority or jurisdiction of one over another, what any may do in prosecution of that law, everyone must needs have a right to do.

And thus in the state of nature one man comes by a power over another; but yet no absolute or arbitrary power, to use a criminal, when he has got him in his hands, according to the passionate heats or boundless extravagance of his own will; but only to retribute to him so far as calm reason and conscience dictate what is proportionate to his transgression, which is so much as may serve for reparation and restraint. For these two are the only reasons why one man may lawfully do harm to another, which is that we call punishment. In transgressing the law of nature, the offender declares himself to live by another rule than that of common reason and equity, which is that measure God has set to the actions of men, for their mutual security; and so he becomes dangerous to mankind, the tie which is to secure them from injury and violence being slighted and broken by him. Which, being a trespass against the whole species, and the peace and safety of it, provided for by the law of nature, every man upon this score, by the right he hath to preserve mankind in general, may restrain, or, where it is necessary, destroy things noxious to them, and so may bring such evil on anyone who hath transgressed that law, as may make him repent the doing of it, and thereby deter him, and by his example others, from doing the like mischief. And in this case, and upon this ground, every man hath a right to punish the offender, and be executioner of the law of nature. . . .

'Tis often asked as a mighty objection, Where are, or ever were there, any men in such a state of nature? To which it may suffice as an answer at present: That since all princes and rulers of independent governments all through the world are in a state of nature, 'tis plain the world never was, nor ever will be, without numbers of men in that state. I have named all governors of independent communities, whether they are or are not in league with others. For 'tis not every compact that puts an end to the state of nature between men, but only this one of agreeing together mutually to enter into one community, and make one body politic; other promises and compacts men may make one with another, and yet still be in the state of nature. The promises and bargains for truck, etc., between the two men in Soldania, in or between a Swiss and an Indian, in the woods of America, are binding to them, though they are perfectly in a state of nature in reference to one another. For truth and keeping of faith belong to men as men, and not as members of society. . . .

OF PROPERTY

Whether we consider natural reason, which tells us that men being once born have a right to their preservation, and consequently to meat and drink and such other things as nature affords for their subsistence; or revelation, which give us an account of those grants God made of the world to Adam, and to Noah and his sons, 'tis very clear that God, as King David says, Pslam cxv. 16, "has given the earth to the children of men," given it to mankind in common. But this being supposed, it seems to some a very great difficulty how anyone should ever come to have a property in anything. I will not content myself to answer that if it be difficult to make out property upon a supposition that God gave the world to Adam and his posterity in common, it is impossible that any man but one universal monarch should have any property upon a supposition that God gave the world to Adam and his heirs in succession, exclusive of all the rest of his posterity. But I shall endeavor to show how men might come to have a property in several parts of that which God gave to mankind in common, and that without any express compact of all the commoners.

God, who hath given the world to men in common, hath also given them reason to make use of it

to the best advantage of life and convenience. The earth and all that is therein is given to men for the support and comfort of their being. And though all the fruits it naturally produces, and beasts it feeds, belong to mankind in common, as they are produced by the spontaneous hand of nature; and nobody has originally a private dominion exclusive of the rest of mankind in any of them as they are thus in their natural state; yet being given for the use of men, there must of necessity be a means to appropriate them some way or other before they can be of any use or at all beneficial to any particular man. The fruit or venison which nourishes the wild Indian, who knows no enclosure, and is still a tenant in common, must be his, and so his, i.e., a part of him, that another can no longer have any right to it, before it can do any good for the support of his life.

Though the earth and all inferior creatures be common to all men, yet every man has a property in his own person; this nobody has any right to but himself. The labor of his body and the work of his hands we may say are properly his. Whatsoever, then, he removes out of the state that nature hath provided and left it in, he hath mixed his labor with, and joined to it something that is his own, and thereby makes it his property. It being by him removed from the common state nature placed it in, it hath by this labor something annexed to it that excludes the common right of other men. For this labor being the unquestionable property of the laborer, no man but he can have a right to what that is once joined to, at least where there is enough, and as good left in common for others.

He that is nourished by the acorns he picked up under an oak, or the apples he gathered from the trees in the wood, has certainly appropriated them to himself. Nobody can deny but the nourishment is his. I ask, then, When did they begin to be his—when he digested, or when he ate, or when he boiled, or when he brought them home, or when he picked them up? And 'tis plain if the first gathering made them not his, nothing else could. That labor put a distinction between them and common; that added something to them more than nature, the common mother of all, had done, and so they became his private right. And will anyone say he had no right to those acorns or apples he thus appropriated, because he had not the consent of all mankind to make them his? Was it robbery thus to assume to himself what belonged to all in common? If such a consent as that was necessary, man had starved, notwithstanding the plenty God had given him. We see in commons which remain so by compact that 'tis the taking any part of what is common and removing it out of the state nature leaves it in, which begins the property; without which the common is of no use. And the taking of this or that part does not depend on the express consent of all the commoners. Thus the grass my horse has bit, the turfs my servant has cut, and the ore I have dug in any place where I have a right to them in common with others, become my property without the assignation or consent of anybody. The labor that was mine removing them out of that common state they were in, hath fixed my property in them. . . .

It will perhaps be objected to this, that if gathering the acorns, or other fruits of the earth, etc., makes a right to them, then anyone may engross as much as he will. To which I answer, Not so. The same law of nature that does by this means give us property, does also bound that property too. "God has given us all things richly" (1 Tim. vi. 17), is the voice of reason confirmed by inspiration. But how far has He given it us? To enjoy. As much as anyone can make use of to any advantage of life before it spoils, so much he may by his labor fix a property in; whatever is beyond this, is more than his share, and belongs to others. Nothing was made by God for man to spoil or destroy. And thus considering the plenty of natural provisions there was a long time in the world and the few spenders, and to how small a part of that provision the industry of one man could extend itself, and engross it to the prejudice of others—especially keeping within the bounds, set by reason, or what might serve for his use—there could be then little room for quarrels or contentions about property so established.

But the chief matter of property being now not the fruits of the earth, and the beasts that subsist on it, but the earth itself, as that which takes in and carries with it all the rest. I think it is plain that property in that, too, is acquired as the former. As much land as a man tills, plants, improves, cultivates, and can use the product of,

so much is his property. He by his labor does as it were enclose it from the common. Nor will it invalidate his right to say, everybody else has an equal title to it; and therefore he cannot appropriate, he cannot enclose, without the consent of all his fellow-commoners, all mankind. God, when He gave the world in common to all mankind, commanded man also to labor, and the penury of his condition required it of him. God and his reason commanded him to subdue the earth, i.e., improve it for the benefit of life, and therein lay out something upon it that was his own, his labor. He that, in obedience to this command of God, subdued, tilled, and sowed any part of it, thereby annexed to it something that was his property, which another had no title to, nor could without injury take from him.

Nor was this appropriation of any parcel of land, by improving it, any prejudice to any other man, since there was still enough and as good left; and more than the yet unprovided could use. So that in effect there was never the less left for others because of his enclosure for himself. For he that leaves as much as another can make use of, does as good as take nothing at all. Nobody could think himself injured by the drinking of another man, though he took a good draught, who had a whole river of the same water left him to quench his thirst; and the case of land and water, where there is enough of both, is perfectly the same. . . .

Before the appropriation of land, he who gathered as much of the wild fruit, killed, caught, or tamed as many of the beasts as he could; he that so employed his pains about any of the spontaneous products of nature as any way to alter them from the state which nature put them in, by placing any of his labor on them, did thereby acquire a propriety in them. But if they perished in his possession without their due use; if the fruits rotted, or the venison putrefied before he could spend it, he offended against the common law of nature, and was liable to be punished; he invaded his neighbor's share, for he had no right further than his use called for any of them and they might serve to afford him conveniences of life.

The same measures governed the possessions of land, too. Whatsoever he tilled and reaped, laid up, and made use of before it spoiled, that was his peculiar right; whatsoever he enclosed and could feed and make use of, the cattle and product was also his. But if either the grass of his enclosure rotted on the ground, or the fruit of his planting perished without gathering and laying up, this part of the earth, notwithstanding his enclosure, was still to be looked on as waste, and might be the possession of any other. . . .

The greatest part of things really useful to the life of man, and such as the necessity of subsisting made the first commoners of the world look after, as it doth the Americans now, are generally things of short duration, such as, if they are not consumed by use, will decay and perish of themselves: gold, silver, and diamonds are things that fancy or agreement have put the value on more than real use and the necessary support of life. Now, of those good things which nature hath provided in common, everyone hath a right, as hath been said, to as much as he could use, and had a property in all he could effect with his labor—all that his industry could extend to, to alter from the state nature had put it in, was his. He that gathered a hundred bushels of acorns or apples had thereby a property in them; they were his goods as soon as gathered. He was only to look that he used them before they spoiled, else he took more than his share, and robbed others; and, indeed, it was a foolish thing, as well as dishonest, to hoard up more than he could make use of. If he gave away a part to anybody else, so that it perished not uselessly in his possession, these he also made use of; and if he also bartered away plums that would have rotted in a week, for nuts that would last good for his eating a whole year, he did no injury; he wasted not the common stock, destroyed no part of the portion of goods that belonged to others, so long as nothing perished uselessly in his hands. Again, if he would give his nuts for a piece of metal, pleased with its color, or exchange his sheep for shells, or wool for a sparkling pebble or a diamond, and keep those by him all his life, he invaded not the right of others; he might heap up as much of these durable things as he pleased, the exceeding of the bounds of his just property not lying in the largeness of his possessions, but the perishing of anything uselessly in it.

And thus came in the use of money—some lasting thing that men might keep without spoiling,

and that, by mutual consent, men would take in exchange for the truly useful but perishable supports of life.

And as different degrees of industry were apt to give men possessions in different proportions, so this invention of money gave them the opportunity to continue and enlarge them; for supposing an island, separate from all possible commerce with the rest of the world, wherein there were but a hundred families—but there were sheep, horses, and cows, with other useful animals, wholesome fruits, and land enough for corn for a hundred thousand times as many, but nothing in the island, either because of its commonness or perishableness, fit to supply the place of money—what reason could anyone have there to enlarge his possessions beyond the use of his family and a plentiful supply to its consumption, either in what their own industry produced, or they could barter for like perishable useful commodities with others? Where there is not something both lasting and scarce, and so valuable to be hoarded up, there men will not be apt to enlarge their possessions of land, were it never so rich, never so free for them to take; for I ask, what would a man value ten thousand or a hundred thousand acres of excellent land, ready cultivated, and well stocked too with cattle, in the middle of the inland parts of America, where he had no hopes of commerce with other parts of the world, to draw money to him by the sale of the product? It would not be worth the enclosing, and we should see him give up again to the wild common of nature whatever was more than would supply the conveniences of life to be had there for him and his family. . . .

But since gold and silver, being little useful to the life of man in proportion to food, raiment, and carriage, has its value only from the consent of men, whereof labor yet makes, in great part, the measure, it is plain that the consent of men have agreed to a disproportionate and unequal possession of the earth—I mean out of the bounds of society and compact; for in governments the laws regulate it; they having, by consent, found out and agreed in a way how a man may rightfully and without injury possess more than he himself can make use of by receiving gold and silver, which may continue long in a man's possession, without decaying for the overplus, and agreeing those metals should have a value.

And thus, I think, it is very easy to conceive without any difficulty how labor could at first begin a title of property in the common things of nature, and how the spending it upon our uses bounded it; so that there could then be no reason of quarrelling about title, nor any doubt about the largeness of possession it gave. Right and conveniency went together; for as a man had a right to all he could employ his labor upon, so he had no temptation to labor for more than he could make use of. This left no room for controversy about the title, nor for encroachment on the right of others; what portion a man carved to himself was easily seen, and it was useless, as well as dishonest, to carve himself too much, or take more than he needed.

REVIEW AND DISCUSSION QUESTIONS

1. Describe the "state of nature" as envisioned by Locke. How is this a state of equality? What does Locke mean when he writes that it is a state of liberty but not of license? What do you think Locke means by the "law of nature"?
2. If the world is given to human beings in common, how can we be justified in appropriating anything for ourselves?
3. Locke states that "every man has a property in his own person." What are the implications of this for Locke?
4. What restrictions does Locke place on the acquisition of resources? Why?
5. What is the origin of money? How does its introduction change things?
6. Has Locke succeeded in justifying private property and inequalities of wealth and possession? Explain why or why not.

Property and Sovereignty

Morris Raphael Cohen

Morris Cohen was one of the leading legal philosophers of the twentieth century. In this essay, be critically examines four justifications of private property: the occupation theory, the labor theory, personal self-assertion, and the economic theory. Although each points to an important value or interest that the legal order needs to protect, private property is only one among other human interests: It is not sacrosanct, and state interference with private property—even the confiscation or abolition of certain types of it—may be justified.

THE JUSTIFICATION OF PROPERTY

1. The Occupation Theory

The oldest and up to recently the most influential defense of private property was based on the assumed right of the original discoverer and occupant to dispose of that which thus became his. This view dominated the thought of Roman jurists and of modern philosophers—from Grotius to Kant—so much so that the right of the laborer to the produce of his work was sometimes defended on the ground that the laborer "occupied" the material which he fashioned into the finished product.

It is rather easy to find fatal flaws in this view. Few accumulations of great wealth were ever simply found. Rather were they acquired by the labor of many, by conquest, by business manipulation, and by other means. It is obvious that today at any rate few economic goods can be acquired by discovery and first occupancy.[1] Even in the few cases when they are, as in fishing and trapping, we are apt rather to think of the labor involved as the proper basis of the property acquired. Indeed, there seems nothing ethically self-evident in the motto that "findings is keepings." There seems nothing wrong in a law that a treasure trove shall belong to the king or the state rather than to the finder. Shall the finder of a river be entitled to all the water in it?

Moreover, even if we were to grant that the original finder or occupier should have possession as against anyone else, it by no means follows that he may use it arbitrarily or that his rule shall prevail indefinitely after his death. The right of others to acquire the property from him by bargain, by inheritance, or by testamentary disposition, is not determined by the principle of occupation.

Despite all these objections, however, there is a kernel of positive value in this principle. Protecting the discoverer or first occupant, is really part of the more general principle that possession as such should be protected. There is real human economy in doing so until somebody shows a better claim than the possessor. It makes for certainty and security of transaction as well as for public peace—provided the law is ready to set aside possession acquired in ways that are inimical to public order. Various principles of justice may determine the distribution of goods and the retribution to be made for acts of injustice. But the law must not ignore the principle of inertia in human affairs. Continued possession creates expectations in the possessor and in others and only a very poor morality would ignore the hardship of frustrating these expectations and rendering human relations insecure, even to correct some old flaws in the original acquisition. Suppose some remote ancestor of yours did acquire your property by fraud, robbery or conquest, *e.g.*, in the days of William of Normandy. Would it be just to take it away from you and your dependents who have held it in good faith? Reflection on the general insecurity that would result from such procedure leads us to see that as habit is the basis of individual life, continued practice must be the basis of social procedure. Any form of property which exists has therefore a claim to continue until it can be shown that the effort to change it is worth while. Continual changes in property laws would certainly discourage enterprise.

Nevertheless, it would be as absurd to argue that the distribution of property must never be modified by law as it would be to argue that the distribution of political power must never be changed. No less a philosopher than Aristotle argued against changing even bad laws, lest the habit of obedience be thereby impaired. There is something to be said for this, but only so long as we are in the realm of merely mechanical obedience. When we introduce the notion of free or rational obedience, Aristotle's argument loses its force in the political realm; and similar considerations apply to any property system that can claim the respect of rational beings.

2. The Labor Theory

That everyone is entitled to the full produce of his labor is assumed as self-evident by both socialists and conservatives who believe that capital is the result of the savings of labor. However, as economic goods are never the result of any one man's unaided labor, our maxim is altogether inapplicable. How shall we determine what part of the value of a table should belong to the carpenter, to the lumberman, to the transport worker, to the policeman who guarded the peace while the work was being done, and to the indefinitely large numbers of others whose co-operation was necessary? Moreover, even if we could tell what any one individual has produced—let us imagine a Robinson Crusoe growing up all alone on an island and in no way indebted to any community—it would still be highly questionable whether he has a right to keep the full produce of his labor when some shipwrecked mariner needs his surplus food to keep from starving.

In actual society no one ever thinks it unjust that a wealthy old bachelor should have part of his presumably just earnings taken away in the form of a tax for the benefit of other people's children, or that one immune to certain diseases, should be taxed to support hospitals, etc. We do not think there is any injustice involved in such cases because social interdependence is so intimate that no man can justly say: "This wealth is entirely and absolutely mine as the result of my own unaided effort."

The degree of social solidarity varies, of course; and it is easy to conceive of a sparsely settled community, such as Missouri at the beginning of the 19th century, where a family of hunters or isolated cultivators of the soil might regard everything which it acquired as the product of its own labor. Generally, however, human beings start with a stock of tools or information acquired from others and they are more or less dependent upon some government for protection against foreign aggression, etc.

Yet despite these and other criticisms, the labor theory contains too much substantial truth to be brushed aside. The essential truth is that labor has to be encouraged and that property must be distributed in such a way as to encourage ever greater efforts at productivity.

As not all things produced are ultimately good, as even good things may be produced at an unjustified expense in human life and worth, it is obvious that other principles besides that of labor or productivity are needed for an adequate basis or justification of any system of property law. We can only say dialectically that all other things being equal, property should be distributed with due regard to the productive needs of the community. We must, however, recognize that a good deal of property accrues to those who are not productive,[2] and a good deal of productivity does not and perhaps should not receive its reward in property. Nor should we leave this theme without recalling the Hebrew-Christian view—and for that matter, the specifically religious view—that the first claim on property is by the man who needs it rather than the man who has created it. Indeed, the only way of justifying the principle of distribution of property according to labor is to show that it serves the larger social need.

The occupation theory has shown us the necessity for security of possession and the labor theory the need for encouraging enterprise. These two needs are mutually dependent. Anything which discourages enterprise makes our possessions less valuable, and it is obvious that it is not worth while engaging in economic enterprise if there is no prospect of securely possessing the fruit of it. Yet there is also a conflict between these two needs. The owners of land, wishing to secure the continued possession by the family, oppose laws which make it subject to free financial transactions or make it possible that land should be taken away from one's heirs by a judgment creditor for personal debts. In an agricultural economy security of possession demands that the owner of a horse

should be able to reclaim it no matter into whose hands it has fallen. But in order that markets should be possible, it becomes necessary that the innocent purchaser should have a good title. . . .

3. Property and Personality

Hegel, Ahrens, Lorimer, and other idealists have tried to deduce the right of property from the individual's right to act as a free personality. To be free one must have a sphere of self-assertion in the external world. One's private property provides such an opportunity.

. . . The notion of personality is too vague to enable us to deduce definite legal consequences by means of it. How, for example, can the principle of personality help us to decide to what extent there shall be private rather than public property in railroads, mines, gas-works, and other public necessities?

Not the extremest communist would deny that in the interest of privacy certain personal belongings, such as are typified by the toothbrush, must be under the dominion of the individual owner, to the absolute exclusion of everyone else. This, however, will not carry us far if we recall that the major effect of property in land, in the machinery of production, in capital goods, etc., is to enable the owner to exclude others from *their necessities*, and thus to compel them to serve him. Ahrens, one of the chief expounders of the personality theory, argues "It is undoubtedly contrary to the right of personality to have persons dependent on others on account of material goods."[3] But if this is so, the primary effect of property on a large scale is to limit freedom, since the one thing that private property law does not do is to guarantee a minimum of subsistence or the necessary tools of freedom to everyone. So far as a regime of private property fails to do the latter it rather compels people to part with their freedom.

It may well be argued in reply that just as restraining traffic rules in the end give us greater freedom of motion, so, by giving control over things to individual property owners, greater economic freedom is in the end assured to all. This is a strong argument, as can be seen by comparing the different degrees of economic freedom that prevail in lawless and in law abiding communities. It is, however, an argument for legal order rather than any particular form of government or private property. It argues for a regime where every one has a definite sphere of rights and duties, but it does not tell us where these lines should be drawn. The principle of freedom of personality certainly cannot justify a legal order wherein a few can, by virtue of their legal monopoly over necessities, compel others to work under degrading and brutalizing conditions. A government which limits the right of large land-holders limits the rights of property and yet may promote real freedom. Property owners, like other individuals, are members of a community and must subordinate their ambition to the larger whole of which they are a part. They may find their compensation in spiritually identifying their good with that of the larger life.

4. The Economic Theory

The economic justification of private property is that by means of it a maximum of productivity is promoted. The classical economic argument may be put thus: "The successful business man, the one who makes the greatest profit, is the one who has the greatest power to foresee effective demand. If he has not that power his enterprise fails. He is therefore, in fact, the best director of economic activities.

There can be little doubt that if we take the whole history of agriculture and industry, or compare the economic output in countries like Russia with that in the United States, there is a strong *prima facie* case for the contention that more intensive cultivation of the soil and greater productiveness of industry prevail under individual ownership. Many *a priori* psychologic and economic reasons can also be brought to explain why this must be so, why the individual cultivator will take greater care not to exhaust the soil, etc. All this, however, is so familiar that we may take it for granted and look at the other side of the case, at the considerations which show that there is a difference between socially desirable productivity and the desire for individual profits.

In the first place let us note that of many things the supply is not increased by making them private property. This is obviously true of land in cities and of other monopoly or limited goods. Private ownership of land does not increase the amount of rainfall, and irrigation works to make the land more fruitful have been carried through by government more

than by private initiative. Nor was the productivity of French or Irish lands reduced when the property of their landlords in rent charges and other incidents of seigniorage was reduced or even abolished. In our own days, we frequently see tobacco, cotton or wheat farmers in distress because they have succeeded in raising too plentiful crops; and manufacturers who are well informed know when greater profit is to be made by a decreased output. Patents for processes which would cheapen the product are often bought up by manufacturers and never used. Durable goods which are more economic to the consumer are very frequently crowded out of the market by shoddier goods which are more profitable to produce because of the larger turnover. Advertising campaigns often persuade people to buy the less economical goods and to pay the cost of the uneconomic advice.

In the second place, there are inherent sources of waste in a regime of private enterprise and free competition. If the biologic analogy of the struggle for existence were taken seriously, we should see that the natural survival of the economically fittest is attended, as in the biologic field, with frightful wastefulness. The elimination of the unsuccessful competitor may be a gain to the survivor but all business failures are losses to the community.

Finally, a regime of private ownership in industry is too apt to sacrifice social interests to immediate monetary profits. This shows itself in speeding up industry to such a pitch that men are exhausted in a relatively few years whereas a slower expenditure of their energy would prolong their useful years. It shows itself in the way in which private ownership enterprise has wasted a good deal of the natural resources of the United States to obtain immediate profits. Even when the directors of a modern industrial enterprise see the uneconomic consequences of immediate profits, the demand of shareholders of immediate dividends,[4] and the ease with which men can desert a business and leave it to others to stand the coming losses, all tend to encourage ultimately wasteful and uneconomic activity. Possibly the best illustration of this is child labor, which by lowering wages increases immediate profits, but in the end is really wasteful of the most precious wealth of the country, its future manhood and womanhood.

Surveying our arguments thus far: We have seen the roots of property in custom and in the need for economic productivity, in individual needs of privacy and in the need for social utility. But we have also noted that property, being only one among other human interests, cannot be pursued absolutely without detriment to human life. Hence we can no longer maintain Montesquieu's view that private property is sacrosanct and that the general government must in no way interfere with or retrench its domain. The issue before thoughtful people is therefore not the maintenance or abolition of private property, but the determination of the precise lines along which private enterprise must be given free scope and where it must be restricted in the interests of the common good.

LIMITATIONS OF PROPERTY RIGHTS

The traditional theory of rights, and the one that still prevails in this country, was molded by the struggle in the 17th and 18th centuries against restrictions on individual enterprise. These restrictions in the interest of special privilege were fortified by the divine (and therefore absolute) rights of kings. As is natural in all revolts, absolute claims on one side were met with absolute denials on the other. Hence the theory of the natural rights of the individual took not only an absolute but a negative form; men have inalienable rights, the state must never interfere with private property, etc. The state, however, must interfere in order that individual rights should become effective and not degenerate into public nuisances. To permit anyone to do absolutely what he likes with his property in creating noise, smells, or danger of fire, would be to make property in general valueless. To be really effective, therefore, the right of property must be supported by restrictions or positive duties on the part of owners, enforced by the state as much as the right to exclude others which is the essence of property. Unfortunately, however, whether because of the general decline of juristic philosophy after Hegel or because law has become more interested in defending property against attacks by socialists, the doctrine of natural rights has remained in the negative state and has never developed into a doctrine of the positive contents of rights.[5] . . .

As a believer in natural rights, I believe that the state can, and unfortunately often does, enact

unjust laws. But I think it is a sheer fallacy based on verbal illusion to think that the rights of the community against an individual owner are no better than the rights of a neighbor. Indeed, no one has in fact had the courage of this confusion to argue that the state has no right to deprive an individual of property to which he is so attached that he refuses any money for it. Though no neighbor has such a right the public interest often justly demands that a proprietor shall part with his ancestral home to which he may be attached by all the roots of his being.

When taking away a man's property, is the state always bound to pay a direct compensation? I submit that while this is generally advisable in order not to disturb the general feeling of security, no absolute principle of justice requires it. I have alluded before to the fact that there is no injustice in taxing an old bachelor to educate the children of others, or to tax one immune to typhoid for the construction of sewers or other sanitary measures. We may go farther and say that the whole business of the state depends upon its rightful power to take away the property of some (in the form of taxation) and use it to support others, such as the needy, those invalided in the service of the state in war or peace, and those who are not yet able to produce but in whom the hope of humanity is embodied. Doubtless, taxation and confiscation may be actuated by malice and may impose needless and cruel hardship on some individuals or classes. But this is not to deny that taxation and confiscation are within the just powers of the state. A number of examples may make this clearer.

A. Slavery. When slavery is abolished by law, the owners have their property taken away. Is the state ethically bound to pay them the full market value of their slaves? It is doubtless a grievous shock to a community to have a large number of slave owners, whose wealth often makes them leaders of culture, suddenly deprived of their income. It may also be conceded that it is not always desirable for the slave himself to be suddenly taken away from his master and cut adrift on the sea of freedom. But when one reads of the horrible ways in which some of those slaves were violently torn from their homes in Africa and shamelessly deprived of their human rights, one is inclined to agree with Emerson that compensation should first be paid to the slaves. This compensation need not be in the form of a direct bounty to them. It may be more effectively paid in the form of rehabilitation and education for freedom; and such a charge may take precedence over the claims of the former owners. After all, the latter claims are no greater than those of a protected industry when the tariff is removed. If the state should decide that certain import duties, *e.g.*, those on scientific instruments, or hospital supplies, are unjustified and proceed to abolish them, many manufacturers may suffer. Are they entitled to compensation by the state?

It is undoubtedly for the general good to obviate as much as possible the effect of economic shock to a large number of people. The routine of life prospers on security. But when that security contains a large element of injustice the shock of an economic operation by law may be necessary and ethically justified.

This will enable us to deal with other types of confiscation.

B. Financial loss through the abolition of public office. It is only in very recent times that we have got into the habit of ignoring the fact that public office is and always has been regarded as a source of revenue like any other occupation. When, therefore, certain public offices are abolished for the sake of good government, a number of people are deprived of their expected income. In the older law and often in popular judgment of today this does not seem fair. But reflection shows that the state is not obligated to pay anyone when it finds that particular services of his are unnecessary. At best, it should help him to find a new occupation.

Part of the prerogative of the English or Scotch landlord was the right to nominate the priest for the parish on his land. To abolish this right of advowson is undoubtedly a confiscation of a definite property right. But . . . I do not see that there could have been any sound ethical objection to the legislature changing the law without compensating the landlord.

C. In our own day, we have seen the confiscation of many millions of dollars of property through prohibition. Were the distillers and brewers entitled to compensation for their losses? We have seen that property on a large scale is power and the loss of it, while evil to those who are accustomed to exercise it, may not be an evil to the community. In point of fact, the shock to the distillers and brewers was

not as serious as to others, *e.g.*, saloon keepers and bartenders who did not lose any legal property since they were only employees, but who found it difficult late in life to enter new employments.

History is full of examples of valuable property privileges abolished without any compensation, *e.g.*, the immunity of nobles from taxation, their rights to hunt over other people's lands, etc. It would be absurd to claim that such legislation was unjust.

These and other examples of justifiable confiscation without compensation are inconsistent with the absolute theory of private property. An adequate theory of private property, however, should enable us to draw the line between justifiable and unjustifiable cases of confiscation. Such a theory I cannot undertake to elaborate on this occasion, though the doctrine of security of possession and avoidance of unnecessary shock seem to me suggestive. I wish however to urge that if the large property owner is viewed, as he ought to be, as a wielder of power over the lives of his fellow citizens, the law should not hesitate to develop a doctrine as to his positive duties in the public interest. The owner of a tenement house in a modern city is in fact a public official and has all sorts of positive duties. He must keep the halls lighted, he must see that the roof does not leak, that there are fire-escape facilities, he must remove tenants guilty of certain public immoralities, etc., and he is compensated by the fees of his tenants which the law is beginning to regulate. Similar is the case of a factory owner. He must install all sorts of safety appliances, hygienic conveniences, see that the workmen are provided with a certain amount of light, air, etc.

In general, there is no reason for the law insisting that people should make the most economic use of their property. They have a motive in doing so themselves and the cost of the enforcing machinery may be a mischievous waste. Yet there may be times, such as occurred during the late war, when the state may insist that man shall cultivate the soil intensively and be otherwise engaged in socially productive work.

With considerations such as these in mind, it becomes clear that there is no unjustifiable taking away of property when railroads are prohibited from posting notice that they will discharge their employees if the latter join trade unions, and that there is no property taken away without due or just process of law when an industry is compelled to pay its laborers a minimum of subsistence instead of having it done by private or public charity or else systematically starving its workers.

NOTES

1. In granting patents, copyrights, etc., the principle of reward for useful work or to encourage productivity seems so much more relevant that the principle of discovery and first occupancy seems to have little force.
2. Economists often claim that the unearned increment is the greatest source of wealth. See Bull. of Am. Econ. Ass'n (4th ser., No. 2) 542 ff.
3. *Il Cours de Droit Naturel* (6th ed.) 108.
4. Thus the leading brewers doubtless foresaw the coming of prohibition and could have saved millions in losses by separating their interests from that of the saloon. But the large temporary loss involved in such an operation was something that stockholders could never have agreed to.
5. Thus our courts are reluctant to admit that rules against unfair competition may be in the interest of the general public and not merely for those whose immediate property interests are directly affected. . . .

REVIEW AND DISCUSSION QUESTIONS

1. Restate each of the four justifications of property discussed by Cohen. What does he see as the strong and weak points of each? Do you agree? What do you see as the most compelling argument for private property?
2. Why in Cohen's view is private property not "sacrosanct"?
3. Would Cohen agree with Locke that we have a natural, pre-political right to property?
4. According to Cohen, does the state have the right to deprive an individual of property? Must it pay compensation? Explain.

Intellectual Property

Richard A. Posner

Built on ideas and technological creativity, the information economy continues to grow in importance. As a result of this growth, recent years have seen an explosion of legal disputes involving intellectual property. With the emergence of new technologies and the "seemingly relentless expansion of intellectual property," the law has, Judge Posner believes, begun to lose its bearing. After reviewing the nature and rationale of property rights in general, he examines the distinguishing (and sometimes paradoxical) characteristics of intellectual property and explains the workings of, and justification for, patents and copyrights. The challenge facing the law is to strike the right balance between society's interest in encouraging the production of intellectual property and its interest in promoting its widespread use.

Legal disputes over intellectual property have exploded in recent years. No field of law is in greater ferment. And in no field of law have judges and scholars experienced more difficulty recently in getting their bearings.

The increase in intellectual property litigation was made inevitable by the rise of the information economy, an economy built on intellectual property—which is now, incidentally, America's largest export. Recognition of the importance of intellectual property in the current American scene is one of the things that lie behind the seemingly relentless expansion of intellectual property rights in modern law.

Two illustrations of that expansion: the copyright term has been repeatedly enlarged in recent years, to the point where copyrights are as a practical matter nearly perpetual. And the new "business method" patents create the potential for inventors of new methods of doing business to obtain enormous monopoly power (imagine if the first person to think up the auction had been able to patent it); such patents also create a reward greatly in excess of the cost of the invention.

The emergence of new technologies has further caused the law to lose its bearings, and this in two respects. First, one of the most important of these new technologies, computer software, is characterized by high monopoly potential conjoined with an extreme disparity between the cost of creation and the cost of making and distributing copies, which indeed approaches zero whenever the copy is made electronically and is distributed over the Internet. Property rights in software may enable its creators to reap enormous profits by charging prices that inhibit distribution, while denying property rights may, in the interest of discouraging excessive investment in software creation and of maximizing distribution, kill the goose that lays the golden eggs by depriving the creators of software of the profit opportunities needed to finance that creation. That is the essential dilemma in crafting a sensible, efficient regime of intellectual property rights.

Second, the products of the new technologies are sometimes hard to fit into the law's pigeonholes. Computer software is a kind of text, which implies that copyright is the proper regime; a kind of machine, which implies that patent is the proper regime; and a kind of algorithm, which traditionally has not been protected by either body of law. In biotechnology, the creation of new forms of life by genetic engineering poses acutely the question of just what should be regarded as patentable technology. . . .

The information economy and its associated novel technologies arose against the background of a mature system of intellectual property law, one that had evolved over centuries out of ancient concepts developed to deal with tangible property. To understand the law of intellectual property, and the muddle we're now in, you must first understand the law of physical property and the social objectives of that law.

When lawyers speak of a "property right," say to a parcel of land, they mean simply that the holder of the right is entitled to invoke the aid of the state to prevent anyone from entering upon the land without his consent. There are all manner of qualifications

of this right (eminent domain, for example—the landowner can't prevent the government from taking his land for a public use, although he can insist that the government pay him just compensation for the taking), but they can be ignored.

What cannot be ignored is why property rights are granted—what social functions they serve. Two are paramount. First, without exclusive rights to the use of tracts of land or other valuable physical objects, these properties would be overused—if anyone has the right to graze his cattle on a pasture, the pasture will be overgrazed and hence depleted prematurely, because each cattle owner will tend to ignore the costs that the grazing by his cattle imposes on the other users of the pasture. Second, without exclusive rights, there will be insufficient incentives to invest in improving property: if you cannot be assured of being able to reap where you have sown, you won't sow, and the land will lie fallow.

It is understood, however, that the social benefits of property rights must be balanced against the costs. When property has little value relative to the costs of creating or enforcing a property right, the right is withheld. Here is a homely example: owners of shopping centers do not charge a price for parking in the shopping center's parking lot. In effect, the owner declines to enforce his property right in the lot, treating it instead as the common property of the shopping center's customers, like a common pasture. This is because charging a fee for entry to the lot, while it would have an economizing effect (the lot could be smaller if access to it were rationed by price, just as tolls limit highway traffic), would cost more than it would be worth; part of the cost would be discouraging people from shopping at the center.

We can follow these themes into the law of intellectual property, but with important qualifications. One is that, in contrast to the grazing example, the use of intellectual property by one person usually doesn't diminish its value to other users. That's because the copies of such property can be multiplied indefinitely at little added cost. If I read a book, I do not deprive others of the use of the intellectual property constituting that book, because they can buy and read other copies without interfering with me. Indeed, widespread use of intellectual property can actually increase the value of the property; in effect, additional copies have negative cost, when the value they confer is taken into account. A popular book or movie becomes a focus of discussion; the more popular it is, the more "copies" of it (in effect) there are, the greater the value.

There is an interesting exception, however, concerning what is called the "right of publicity," confusingly classified as part of the "right of privacy." A person has a right not to have his name or likeness used for advertising or other commercial purposes without his consent. This is a right particularly valued by celebrities. Should there be such a right? Does it have useful incentive effects, comparable to the effect of granting property rights to land to create the incentive to cultivate the land? And even if it does, what should happen when the celebrity dies? Should the right die with the celebrity, on the theory that he will no longer be "incentivized" by it to cultivate his image and that therefore anyone should be free to use his name and likeness in advertising? The answer is No, for the same reason that property rights are recognized even in "natural" pastures, that is, pastures not created or improved by investing, unlike ordinary farmland: there would be overuse. The advertising value of the celebrity would be reduced if the celebrity's name and likeness could be attached to an indefinite number of different products. There can be "congestion" even of intellectual property. And this is true whether or not the celebrity is still alive.

Still, in general, the use of intellectual property by one person does not reduce the value of its use by another. Stated differently, the marginal cost of intellectual property—the cost of adding one more user of it—is very low. As I noted earlier, it is essentially zero in the case of computer software, which can be delivered to a new user over the Internet—and it can even be negative.

This has led some students of intellectual property to think it would be desirable to make such property available for free to anyone who wanted to use it, since, in general, optimum output is achieved by equating price to marginal cost, and in the case of much intellectual property this means setting the price at (or only trivially above) zero—or even subsidizing distribution.

But as is now well understood, such a policy would be disastrous. It would kill the incentive to create the intellectual property in the first place, outside of the relatively rare cases in which the creators have powerful nonmonetary incentives to create such property, or in which its creation is

financed other than by sale or lease of the property (by taxation, for example, or charitable donation—such as the patronage of authors by wealthy people, in the old days). We need not suppose that most creative people are greedy to realize that if they cannot obtain a pecuniary benefit from producing intellectual property they will not be able to finance the costs (including the costs of their time) required to produce it.

And so the state defines, recognizes, and enforces property rights in intellectual property. The most important such rights are copyrights and patents, the former a property right in expression, the latter a property right in useful ideas. . . .

Copyrights and patents are both limited in duration, unlike rights in physical property, and an initial question is why? There are several answers, and they point to the fundamental differences between physical and intellectual property.

One answer is the tracing problem, which looms large in the definition and enforcement of intellectual property generally. Items of physical property are visibly distinct; this is true even of adjacent parcels of land, once the boundary has been mapped and fenced. But one piece of intellectual property is not visibly distinct from others; it is identified only by comparison with others. Two copies of the same book are physically distinct, but the intellectual property contained in them is identical. Worse, two different books may be sufficiently similar to raise a question of whether the intellectual property in one was appropriated by the author of the other. If copyright were perpetual, James Joyce or his publisher would have become embroiled in litigation with the heirs of Homer over whether *Ulysses* infringed the *Odyssey*, and Leonard Bernstein with the heirs of Ovid over whether *West Side Story* infringed *Pyramus and Thisbe* (not to mention *Romeo and Juliet* and *A Midsummer Night's Dream*, themselves arguably infringements of Ovid's story). If patents were perpetual, heirs of Leonardo da Vinci would be litigating over rights to basic aircraft technology.

The tracing problem is more serious for copyrights than for patents. The Patent and Trademark Office contains descriptions of patents classified by subject matter, and it is feasible though often difficult to search through these descriptions and identify the patents that a proposed new patent may infringe. But it is impossible as yet to search through the entire body of copyrighted materials. That is one reason why copyright protection is more limited than patent protection. A copyright is infringed only if it is copied; if it is duplicated innocently, there is no infringement. A patent is infringed by being duplicated, even if the duplication was innocent—a case of independent discovery.

Even in the case of copyright, however, the tracing problem is really rather superficial. If copyright owners were required to renew their copyrights periodically by filing a notice of renewal in a public registry, it would be simple enough for creators of new intellectual property to determine whether a given work was in the public domain.

There is a more serious concern with giving the owner of intellectual property too expansive a right. If copyright were perpetual, Ovid's heirs would probably win their suit against Leonard Bernstein; Shakespeare's heirs certainly would (*West Side Story* is based directly on *Romeo and Juliet*)—except they might lose in turn to Ovid's heirs! This means that cutting off copyright protection after a specified period shorter than eternity not only limits tracing costs, but also reduces the pecuniary gain to the owner of the copyright.

There are two reasons why that might be a good thing. First, intellectual property presents a more serious problem of what economists call "rent seeking" than physical property does. A "rent," in economics, is not a rental; it is an excess of revenue over cost. It is pure profit, which is to say profit in excess of the cost of capital (which is not "profit" in an economic sense but merely another cost of doing business). Rent seeking can be bad from a social standpoint because it can lead to excessive investment.

An example is a hunt for buried treasure. If the treasure has a value of $10 million, which will be awarded to the first finder, there will be a race to be first that may eat up the entire profit. Suppose that the cost to a particular finder of finding the treasure by April 1, 2002, would be $1 million. Would-be finders might incur much greater costs in vying to find it sooner—for example, a finder who was confident that by expending an additional $8 million he could win the race by finding the treasure on March 31 would consider the expenditure worthwhile, since it would yield him a profit of $1 million. But the additional cost incurred to win

the race would be wasted from a social standpoint, because the social benefit of finding the treasure a day earlier would be negligible.

The problem of rent seeking is no longer acute in the case of the historically most important form of property, land, because virtually all land is owned. (The situation was quite otherwise in the age of exploration and discovery of new continents.) There would be no rent-seeking problem in the buried-treasure example if someone owned the treasure and were merely offering a reward to the finder—the owner would set the reward at a level designed to obtain the finding service at least cost.

But, as noted, the problem of rent seeking is acute in the "land grab" phase of development—and that is the phase we're perpetually in with regard to intellectual property. For remember that intellectual property is created rather than found, which means that if rights to intellectual property are defined too broadly, the rents generated by them will be so great that excessive resources will be drawn into efforts to be the first to create a valuable piece of intellectual property and thus to obtain the property right to it. Limiting the duration of the property right is one way of cutting down its value to the owner and thus reducing the amount of rent seeking.

But, once again, this concern must not be exaggerated. Because of discounting to present value (that is, the preference for money now over the same sum of money years or decades hence), the difference in value to the creator of intellectual property of, say, a seventy-five-year term and a thousand-year term would actually be very slight, because the present value of a dollar not to be received for seventy-five years (or one hundred or one thousand years) is very slight.

A second reason for wanting to limit the potential reward to owners of intellectual property rights is the previously noted effect of those rights in limiting the distribution and hence use of intellectual property. The fees that the owner of intellectual property charges for its use deflect some users to other products that may cost society more to produce (remember that the marginal cost, the cost of adding one more user, of intellectual property is often close to, at, or even below zero), resulting in a loss of efficiency. Some of those users, moreover, may be other creators of intellectual property, so that expansive intellectual property rights may actually reduce the creation of intellectual property—an important and counterintuitive point to which I'll return.

Against all this must be weighed the incentive effects of allowing the property owner to obtain revenue from property that may have cost him a great deal to create. But it doesn't follow that he has to be able to collect fees *in perpetuity* in order to recoup his investment. Perpetual fees may result in a reward that exceeds the cost of creating the property in the first place, thus resulting in a needless restriction of the use of the property along with the wasteful expenditures caused by rent seeking.

It is true, as I have said, that because most people discount future income steeply, the excess reward that perpetual fees would confer on creators of intellectual property is somewhat illusory. Few people will work harder today to generate some additional income to their heirs (if any) a century hence. But this means that perpetual fees have very little upside in creating incentives for the creation of intellectual property; the tracing costs, and the effect of perpetual copyright in complicating the use of existing intellectual property as an input into new intellectual property, become decisive objections to perpetual rights.

Disregarded in this analysis, however, is the point made earlier in connection with the right of publicity—the potential congestion cost if valuable property is unowned. For example, if anyone can use the character of Mickey Mouse, the public may become tired of him, and his value may drop to zero. Suppose, moreover, that to create a demand for an old expressive work requires a current investment. What publisher would incur the expense and risk of developing a demand for an eighteenth-century author whose works were long out of copyright if the publisher acquired no property right in the works, so that if his expenditures succeeded in creating a demand for them, any other publisher could publish the works without incurring the expense that he had incurred? In both the Mickey Mouse case and in this case, there is overuse because of lack of property rights, but in the first case it leads to the value of the intellectual property plummeting, and in the second case it impairs the incentive to invest in intellectual property.

The solution might be a system of indefinitely renewable copyrights. The initial grant might be for twenty-five years, renewable thereafter every

five years. A stiff fee would assure that most works returned to the public domain. But those works requiring continuing investment or careful management to avoid consumer exhaustion would continue to be owned property.

Copyrights and patents are limited in other ways besides duration. The copyright owner is permitted to copyright only the expressive dimension of the work and not the basic ideas or motifs. Even if copyright were perpetual, Homer's heirs could not demand a royalty for every epic poem written, since the idea of the epic poem (or of rhyme or particular rhyme schemes, or of a story of a war to recover an abducted beauty) would be considered to fall on the idea side of the idea/expression divide. Similarly, patents are limited to ideas that are useful (in the sense of practical, utilitarian), novel, and nonobvious, and so are not available for the fruits of basic research, such as Euclidean geometry, Planck's constant, or $e = mc^2$.

If basic scientific findings were patentable, the tracing problem would be particularly acute. Even more important, patents on basic research would sometimes generate grossly excessive revenues, relative to costs, which in the case of much basic research are low.

Similarly, if valuable applications of scientific theory (as distinct from basic research)—"inventions" or new technology—could be patented in perpetuity, one untoward result would be to limit the use of inventions, and another might be to draw excessive investment into innovation. Bear in mind that the patent process, like my hypothetical hunt for buried treasure, is a race. Whoever crosses the finish line first, if only by a day, receives the entire value of the patent, not the value of accelerating the invention by one day. So we want to make sure that the rewards of owning a patent are not so huge that they operate to suck a disproportionate fraction of society's scarce resources into efforts to accelerate the pace of invention.

As for allowing basic ideas, themes, motifs, character types, and so forth to be copyrighted, the effect in increasing the incentives to create new literature, art, and entertainment by increasing the financial rewards would be more than offset by the effect in discouraging that creation by forcing every new writer to negotiate permission with the heirs of long-deceased predecessors. Literature, art, and entertainment to a great extent play variations on a rather simple, stock set of themes, plots, character types, and so forth. The distinction, the quality, of creative expression lies precisely in the variations, and we want to encourage these by permitting the creators to draw freely on the stock.

A complication is created by the merger of "idea" and "expression" in some forms of modern art, such as Andy Warhol's *Brillo Box*, a work of art that is such not by virtue of any novel or distinctive expression—it is indistinguishable from an ordinary box of Brillo—but solely by virtue of its being treated as art by collectors and museums. In effect, this kind of art is simply the idea of treating an everyday object as a work of art.

I have thus far depicted the basic challenge in the fine-tuning of intellectual property rights as striking the right balance between the interest in encouraging the production of intellectual property and the interest in promoting its widespread use, though I have noted some other concerns as well (such as overinvestment and tracing costs).

But one of the most interesting characteristics of intellectual property, which differentiates it sharply from physical property, is that—paradoxically—limiting intellectual property rights may often be necessary to maximize the creation of intellectual property—in which event the conflict between the creation interest and the use interest disappears. I have given examples of this important point already. Consider now the "fair use" doctrine of copyright law, which permits in specified circumstances some copying of a copyrighted work without having to obtain the owner's consent. An example is quoting from and summarizing a copyrighted book in a review of the book. Suppose such copying required the consent of the book's author or publisher. Then book reviews would lack credibility, since readers would know that the reviewer had a strong incentive to review the book favorably lest publishers refuse to consent to his quoting from subsequent books, or charge him an exorbitant fee for permission to quote. Publishers and authors as a group (though maybe not the publishers and authors of the worst books) would be hurt by a system that deprived readers of the information contained in reviews by people not beholden to the publisher. The publishing industry would lose the most credible form of advertising of its wares.

Similarly, but more fundamentally, anyone familiar with the practices of both authors and inventors knows that most intellectual property, even of a distinctly innovative sort, builds heavily on previous intellectual property (*Ulysses* is again an example). The existing stock of ideas and expression is, to a great extent, the raw material from which new intellectual property is fashioned.

The cheaper a producer's raw materials, the cheaper the final product and so the greater the output. If Joyce had had to negotiate with Homer's heirs over the use of material from the *Odyssey* in his book, it would have taken him longer to write the book: if negotiations had broken down, he might not have been able to write it at all.

. We want, therefore, a process by which intellectual property, having been legally protected in order to create the proper incentives, can eventually be returned to the public domain, there to be available as cheap raw material for future creators of intellectual property. This is another important reason for limiting both the duration of intellectual property rights and their scope.

The economic analysis sketched in this paper is simple, largely intuitive, commonsensical, and, I venture to suggest, fairly uncontroversial. To summarize it, granting property rights in intellectual property increases the incentive to create such property, but the downside is that those rights can interfere with the creation of subsequent intellectual property (because of the tracing problem and because the principal input into most intellectual property is previously created intellectual property). Property rights can limit the distribution of intellectual property and can draw excessive resources into the creation of intellectual property, and away from other socially valuable activities, by the phenomenon of rent seeking.

Striking the right balance, which is to say determining the optimal scope of intellectual property rights, requires a comparison of these benefits and costs—and really, it seems to me, nothing more. The problems are not conceptual; the concepts are straightforward. The problems are entirely empirical. They are problems of measurement.

In addition, we do not know how much intellectual property is in fact socially useful, and therefore we do not know how extensive a set of intellectual property rights we should create. For all we know, too many resources are being sucked into the creation of new biotechnology, computer software, films, pharmaceuticals, and business methods because the rights to these different forms of intellectual property have been too broadly defined.

Unfortunately, the empirical problems are acute—and little progress has been made as yet toward their solution. We urgently need more empirical evidence. The task is daunting, for it requires that we be able to estimate both the social gains from additional intellectual property of different types and the social costs of trying to induce the creation of the additional intellectual property by means of adjustments in the regime of intellectual property rights.

REVIEW AND DISCUSSION QUESTIONS

1. According to Posner, property rights serve two particularly important social functions. What are they? How is intellectual property an exception to the first of these? What about the second?
2. What is the "tracing" problem? What is "rent seeking," and why is it a problem with respect to intellectual property?
3. Posner writes that "limiting intellectual property rights may often be necessary to maximize the creation of intellectual property." Explain. What are the costs and benefits of limiting the potential reward to owners of intellectual property rights?
4. Has American law expanded intellectual property rights too far? For how long a period should copyrights and patents be granted?
5. The owners of some intellectual discoveries believe their property rights should be perpetual. Explain why you agree or disagree. What would Locke say?
6. Is Posner's position compatible with the belief that some property rights (either physical or intellectual) are natural or pre-legal?

Property Acquisition

Haslem v. Lockwood

What constitutes property and under what conditions can one acquire it through labor or possession? These and related questions are addressed in this case. The Court of Common Pleas for Fairfield County, Connecticut, ruled that the plaintiff in this case had not established a sufficient interest in, or right of possession to, the manure he had collected alongside a public road, which the defendant found the next morning and took for himself. On appeal, Connecticut's Supreme Court of Errors overturned this judgment and ordered a new trial.

TROVER [a suit for return of property], for a quantity of manure; brought before a justice of the peace and appealed by the defendant to the Court of Common Pleas for the county of Fairfield, and tried in that court, on the general issue closed to the court, before BREWSTER, J.: On the trial it was proved that the plaintiff employed two men to gather into heaps, on the evening of April 6, 1869, some manure that lay scattered along the side of a public highway, for several rods, in the borough of Stamford, intending to remove the same to his own land the next evening. The men began to scrape the manure into heaps at six o'clock in the evening, and after gathering eighteen heaps, or about six cartloads, left the same at eight o'clock in the evening in the street. The heaps consisted chiefly of manure made by horses hitched to the railing of the public park in, and belonging to the borough of Stamford, and was all gathered between the center of the highway and the park; the rest of the heaps consisting of dirt, straw and the ordinary scrapings of highways. The defendant on the next morning, seeing the heaps, endeavored without success to ascertain who had made them, and inquired of the warden of the borough if he had given permission to any one to remove them, and ascertained from him that he had not. He thereupon, before noon on that day, removed the heaps and also the rest of the manure scattered along the side of the highway adjacent to the park, to his own land.

The plaintiff and defendant both claimed to have received authority from the warden to remove the manure before the 6th of April, but in fact neither had any legal authority from the warden, or from any officer of the borough or of the town. The borough of Stamford was the sole adjoining proprietor of the land on which the manure lay scattered before it was gathered by the plaintiff. No notice was left on the heaps or near by, by the plaintiff or his workmen, to indicate who had gathered them, nor had the plaintiff or his workmen any actual possession of the heaps after eight o'clock in the evening on the 6th of April.

Neither the plaintiff while gathering, nor the defendant while removing the heaps, was interfered with or opposed by any one. The removal of the manure and scrapings was calculated to improve the appearance and health of the borough. The six loads were worth one dollar per load. The plaintiff, on ascertaining that the defendant had removed the manure, demanded payment for the same, which the defendant refused. Neither the plaintiff nor the defendant owned any land adjacent to the place where the manure lay. The highway was kept in repair by the town of Stamford. . . .

The court ruled adversely to the claims of the plaintiff and held that on the facts proved the plaintiff had not made out a sufficient interest in, or right of possession to, the subject matter in dispute, to authorize a recovery in the suit, and rendered judgment for the defendant.

The plaintiff moved for a new trial for error in this ruling of the court.

---

37 Conn. 500 (1871). Citations and notes omitted.

Curtis and Joyt [Counsel for the plaintiff-appellant], in support of the motion:

1. The manure in question was personal property abandoned by its owners.

2. It never became a part of the real estate on which it was abandoned.

3. It being personal property abandoned by its owners, and lying upon the highway, and neither the owners of the fee nor the proper authorities of the town and borough having by any act of theirs shown any intention to appropriate the same, it became lawful for the plaintiff to gather it up and remove it from the highway, providing he did not commit a trespass, and removed it without objection from the owners of the land. And no trespass was in fact committed. No person interfered with the plaintiff or made any objection. This court cannot presume a trespass to have been committed.

4. But if the manure had become a part of the real estate, yet when it was gathered into heaps by the plaintiff it was severed from the realty and became personal estate. And being gathered without molestation from any person owning or claiming to own the land, it is to be considered as having been taken by the tacit consent of such owner.

5. The plaintiff therefore acquired not only a valid legal possession, but a title by occupancy, and by having expended labor and money upon the property. Such a title is a good legal title against every person but the true owner.

6. If the plaintiff had a legal title then he had the constructive possession. If he had legal possession, and only left the property for a short time intending to return and take it away, then he might maintain an action against a wrongdoer for taking it away. The leaving of property for a short time, intending to return, does not constitute an abandonment. The property is still to be considered as in the possession of the plaintiff.

Olmstead [Counsel for the defendant-respondent], contra:

1. The manure mixed with the dirt and ordinary scrapings of the highway, being spread out over the surface of the highway, was a part of the real estate, and belonged to the owner of the fee, subject to the public easement.

2. The scraping up of the manure and dirt into piles, if the same was a part of the real estate, did not change its nature to that of *personal property*, unless there was a severance of it from the realty by removal (which there was not) or whether the plaintiff had the consent of the owner of the fee (which consent it is conceded the plaintiff did not have.)

3. Unless the scraping up of the heaps made their substance *personal property*, the plaintiff could not maintain his action either for trespass or trespass on the case.

4. In trespass *de bonis asportatis*, or Trover, the plaintiff must have had the *actual possession*, or a right to the immediate possession, in order to recover.

5. If the manure was always personal estate, it being spread upon the surface of the earth, it was in possession of the owner of the fee, who was not the plaintiff. The scraping of it into heaps, unless it was removed, would not change the *possession* from the owner of the fee to the plaintiff. The plaintiff therefore never had the *possession.*

6. If the heaps were personal property the plaintiff never had any right in the property, but only *mere possession*, if anything, which he abandoned by leaving the same upon the public highway from 8 o'clock in the evening until 12 o'clock the next day, without leaving any notice on or about the property, or any one to exercise control over the same in his behalf.

Judge Park: We think the manure scattered upon the ground, under the circumstances of this case, was personal property. The cases referred to by the defendant to show that it was real estate are not in point. The principle of those cases is, that manure made in the usual course of husbandry upon a farm is so attached to and connected with the realty that, in the absence of any express stipulation to the contrary, it becomes appurtenant to it. The principle was established for the benefit of agriculture. It found its origin in the fact that it is essential to the successful cultivation of a farm that the manure, produced from the droppings of cattle and swine fed upon the products of the farm, and composted with earth and vegetable matter taken from the land, should be used to supply the drain made upon the soil in the production of crops,

which otherwise would become impoverished and barren; and in the fact that manure so produced is generally regarded by farmers in this country as a part of the realty and has been so treated by landlords and tenants from time immemorial.

But this principle does not apply to the droppings of animals driven by travelers upon the highway. The highway is not used, and cannot be used, for the purpose of agriculture. The manure is of no benefit whatsoever to it, but on the contrary is a detriment; and in cities and large villages it becomes a nuisance, and is removed by public officers at public expense. The finding in this case is, "that the removal of the manure and scrapings was calculated to improve the appearance and health of the borough." It is therefore evident that the cases relied upon by the defendant have no application to the case.

But it is said that if the manure was personal property, it was in the possession of the owner of the fee, and the scraping it into heaps by the plaintiff did not change the possession, but it continued as before, and that therefore the plaintiff cannot recover, for he neither had the possession nor the right to the immediate possession.

The manure originally belonged to the travelers whose animals dropped it, but it being worthless to them was immediately abandoned; and whether it then became the property of the borough of Stamford which owned the fee of the land on which the manure lay, it is unnecessary to determine; for, if it did, the case finds that the removal of the filth would be an improvement to the borough, and no objection was made by any one to the use that the plaintiff attempted to make of it. Considering the character of such accumulations upon highways in cities and villages, and the light in which they are everywhere regarded in closely settled communities, we cannot believe that the borough in this instance would have had any objection to the act of the plaintiff in removing a nuisance that affected the public health and the appearance of the streets. At all events, we think the facts of the case show a significant right in the plaintiff to the immediate possession of the property as against a mere wrongdoer.

The defendant appears before the court in no enviable light. He does not pretend that he had a right to the manure, even when scattered upon the highway, superior to that of the plaintiff; but after the plaintiff had changed its original condition and greatly enhanced its value by his labor, he seized and appropriated to his own use the fruits of the plaintiff's outlay, and now seeks immunity from responsibility on the ground that the plaintiff was a wrong doer as well as himself. The conduct of the defendant is in keeping with his claim, and neither commends itself to the favorable consideration of the court. The plaintiff had the peaceable and quiet possession of the property; and we deem this sufficient until the borough of Stamford shall make complaint.

It is further claimed that if the plaintiff had a right to the property by virtue of occupancy, he lost the right when he ceased to retain the actual possession of the manure after scraping it into heaps.

We do not question the general doctrine, that where the right by occupancy exists, it exists no longer than the party retains the actual possession of the property, or till he appropriates it to his own use by removing it to some other place. If he leaves the property at the place where it was discovered, and does nothing whatsoever to enhance its value or change its nature, his right by occupancy is unquestionably gone. But the question is, if a party finds property comparatively worthless, as the plaintiff found the property in question, owing to its scattered condition upon the highway, and greatly increases its value by his labor and expense, does he lose his right if he leaves it a reasonable time to procure the means to take it away, when such means are necessary for its removal?

Suppose a teamster with a load of grain, while traveling the highway, discovers a rent in one of his bags, and finds that his grain is scattered upon the road for the distance of a mile. He considers the labor of collecting his corn of more value than the property itself, and he therefore abandons it, and pursues his way. *A* afterwards finds the grain in this condition and gathers it kernel by kernel into heaps by the side of the road, and leaves it a reasonable time to procure the means necessary for its removal. While he is gone for his bag, *B* discovers the grain thus conveniently collected in heaps and appropriates it to his own use. Has *A* any remedy? If he has not, the law in this instance is open to just reproach. We think under such circumstances, *A* would

have a reasonable time to remove the property, and during such reasonable time his right to it would be protected. If this is so, then the principle applies to the case under consideration.

A reasonable time for the removal of this manure had not elapsed when the defendant seized and converted it to his own use. The statute regulating the rights of parties in the gathering of sea-weed, gives the party who heaps it upon a public beach twenty-four hours in which to remove it, and that length of time for the removal of the property we think would not be unreasonable in most cases like the present one.

We therefore advise the Court of Common Pleas to grant a new trial.

In this opinion the other judges concurred.

REVIEW AND DISCUSSION QUESTIONS

1. What are the main arguments presented by each of the lawyers on behalf of their clients?
2. Do you agree with the conclusion of the court? Was Judge Park correct to reverse the decision of the trial court? How would you have decided this case?
3. Was the manure private property before being collected by the plaintiff? If so, whose was it? If not, did it become the plaintiff's private property upon being collected? If so, was it still his property when the defendant happened upon it the next day? Why or why not?
4. Will one's thinking about this case be influenced by whether one leans toward a Lockean view of property or toward a perspective more like that of Cohen or Posner? If so, how? If not, explain why not.

Regulation of Private Property

Pennsylvania Coal Co. v. Mahon

The Pennsylvania Coal Company deeded land to Mahon, with the express provision that the company be allowed to mine the coal under that land without responsibility for any damage that might arise from its doing so. The dispute arose when Mahon tried to prevent Pennsylvania Coal from mining under the land because of a Pennsylvania law forbidding the mining of coal in a way that causes the land above it to subside (unless that land is owned by those mining the coal). The case eventually went to the Supreme Court, where the constitutionality of the Pennsylvania statute was the central issue: Does it violate the Fifth Amendment's requirement that "private property [shall not] be taken for public use, without just compensation"?

Mr. Justice Holmes delivered the opinion of the Court: This is a bill in equity brought . . . to prevent the Pennsylvania Coal Company from mining under [the Mahon] property in such way as to remove the supports and cause a subsidence of the surface and of their house. The bill sets out a deed executed by the Coal Company in 1878, under which the plaintiffs claim. The deed conveys the surface, but in express terms reserves the right to remove all the coal under the same, and the grantee takes the premises with the risk, and waives all claim for damages that may arise from mining out the coal. But the plaintiffs say that whatever may have been the Coal Company's rights, they were taken away by an Act of Pennsylvania, approved May 27, 1921, P.L. 1198, commonly known there as the Kohler Act. . . .

The statute forbids the mining of anthracite coal in such way as to cause the subsidence of, among other things, any structure used as a human habitation, with certain exceptions, including among them land where the surface is owned by

260 U.S. 393 (1922).

the owner of the underlying coal and is distant more than one hundred and fifty feet from any improved property belonging to any other person. As applied to this case the statute is admitted to destroy previously existing rights of property and contract. The question is whether the police power can be stretched so far.

Government hardly could go on if to some extent values incident to property could not be diminished without paying for every such change in the general law. As long recognized, some values are enjoyed under an implied limitation and must yield to the police power. But obviously the implied limitation must have its limits, or the contract and due process clauses are gone. One fact for consideration in determining such limits is the extent of the diminution. When it reaches a certain magnitude, in most if not in all cases there must be an exercise of eminent domain and compensation to sustain the act. So the question depends upon the particular facts. The greatest weight is given to the judgment of the legislature, but it always is open to interested parties to contend that the legislature has gone beyond its constitutional power.

This is the case of a single private house. No doubt there is a public interest even in this, as there is in every purchase and sale and in all that happens within the commonwealth. Some existing rights may be modified even in such a case. *Rideout v. Knox*, 148 Mass. 368. But usually in ordinary private affairs the public interest does not warrant much of this kind of interference. A source of damage to such a house is not a public nuisance even if similar damage is inflicted on others in different places. The damage is not common or public. *Wesson v. Washburn Iron Co.*, 13 Allen 95, 103. The extent of the public interest is shown by the statute to be limited, since the statute ordinarily does not apply to land when the surface is owned by the owner of the coal. Furthermore, it is not justified as a protection of personal safety. That could be provided for by notice. Indeed the very foundation of this bill is that the defendant gave timely notice of its intent to mine under the house. On the other hand the extent of the taking is great. It purports to abolish what is recognized in Pennsylvania as an estate in land—a very valuable estate—and what is declared by the Court below to be a contract hitherto binding the plaintiffs. If we were called upon to deal with the plaintiffs' position alone, we should think it clear that the statute does not disclose a public interest sufficient to warrant so extensive a destruction of the defendant's constitutionally protected rights.

But the case has been treated as one in which the general validity of the act should be discussed. . . .

It is our opinion that the act cannot be sustained as an exercise of the police power, so far as it affects the mining of coal under streets or cities in places where the right to mine such coal has been reserved. As said in a Pennsylvania case, "For practical purposes, the right to coal consists in the right to mine it." *Commonwealth v. Clearview Coal Co.*, 256 Pa. St. 328, 331. What makes the right to mine coal valuable is that it can be exercised with profit. To make it commercially impracticable to mine certain coal has very nearly the same effect for constitutional purposes as appropriating or destroying it. This we think that we are warranted in assuming that the statute does.

. . . The protection of private property in the Fifth Amendment presupposes that it is wanted for public use, but provides that it shall not be taken for such use without compensation. . . . When this seemingly absolute protection is found to be qualified by the police power, the natural tendency of human nature is to extend the qualification more and more until at last private property disappears. But that cannot be accomplished in this way under the Constitution of the United States.

The general rule at least is, that while property may be regulated to a certain extent, if regulation goes too far it will be recognized as a taking. It may be doubted how far exceptional cases, like the blowing up a house to stop a conflagration, go—and if they go beyond the general rule, whether they do not stand as much upon tradition as upon principle. . . . In general it is not plain that a man's misfortunes or necessities will justify his shifting the damages to his neighbor's shoulders. We are in danger of forgetting that a strong public desire to improve the public condition is not enough to warrant achieving the desire by a shorter cut than the constitutional way of paying for the change. As we already have said, this is a question of degree—and therefore cannot be disposed of by general propositions. But we regard this as going beyond any of the cases decided by this Court. . . .

We assume, of course, that the statute was passed upon the conviction that an exigency existed that would warrant it, and we assume that an exigency exists that would warrant the exercise of eminent domain. But the question at bottom is upon whom the loss of the changes desired should fall. So far as private persons or communities have seen fit to take the risk of acquiring only surface rights, we cannot see that the fact that their risk has become a danger warrants the giving to them greater rights than they bought.

Decree reversed.

Mr. Justice Brandeis, dissenting: The Kohler Act prohibits, under certain conditions, the mining of anthracite coal within the limits of a city in such a manner or to such an extent "as to cause the . . . subsidence of any dwelling or other structure used as a human habitation, or any factory, store, or other industrial or mercantile establishment in which human labor is employed." Coal in place is land; and the right of the owner to use his land is not absolute. He may not so use it as to create a public nuisance; and uses, once harmless, may, owing to changed conditions, seriously threaten the public welfare. Whenever they do, the legislature has power to prohibit such uses without paying compensation; and the power to prohibit extends alike to the manner, the character and the purpose of the use. Are we justified in declaring that the Legislature of Pennsylvania has, in restricting the right to mine anthracite, exercised this power so arbitrarily as to violate the Fourteenth Amendment? [The Fourteenth Amendment requires, in effect, that individual states like Pennsylvania respect the Bill of Rights, including the Fifth Amendment.—Eds.]

Every restriction upon the use of property imposed in the exercise of the police power deprives the owner of some right theretofore enjoyed, and is, in that sense, an abridgment by the State of rights in property without making compensation. But restriction imposed to protect the public health, safety or morals from dangers threatened is not a taking. The restriction here in question is merely the prohibition of a noxious use. The property so restricted remains in the possession of its owner. The State does not appropriate it or make any use of it. The State merely prevents the owner from making a use which interferes with paramount rights of the public. Whenever the use prohibited ceases to be noxious—as it may because of further change in local or social conditions—the restriction will have to be removed and the owner will again be free to enjoy his property as heretofore.

The restriction upon the use of this property can not, of course, be lawfully imposed, unless its purpose is to protect the public. But the purpose of a restriction does not cease to be public, because incidentally some private persons may thereby receive gratuitously valuable special benefits. . . . Furthermore, a restriction, though imposed for a public purpose, will not be lawful, unless the restriction is an appropriate means to the public end. But to keep coal in place is surely an appropriate means of preventing subsidence of the surface; and ordinarily it is the only available means. Restriction upon use does not become inappropriate as a means, merely because it deprives the owner of the only use to which the property can then be profitably put. . . . Nor is a restriction imposed through exercise of the police power inappropriate as a means, merely because the same end might be effected through exercise of the power of eminent domain, or otherwise at public expense. . . . If by mining anthracite coal the owner would necessarily unloose poisonous gases, I suppose no one would doubt the power of the State to prevent the mining, without buying his coal fields. And why may not the State, likewise, without paying compensation, prohibit one from digging so deep or excavating so near the surface, as to expose the community to like dangers? In the latter case, as in the former, carrying on the business would be a public nuisance.

It is said that one fact for consideration in determining whether the limits of the police power have been exceeded is the extent of the resulting diminution in value; and that here the restriction destroys existing rights of property and contract. But values are relative. If we are to consider the value of the coal kept in place by the restriction, we should compare it with the value of all other parts of the land. That is, with the value not of the coal alone, but with the value of the whole property. The rights of an owner as against the public are not increased by dividing the interests in his property into surface and subsoil. The sum of the rights in the parts can not be greater than the rights in the whole. The

estate of an owner in land is grandiloquently described as extending *ab orco usque ad coelum* ["from the center of the earth to the heavens"—EDS.]. But I suppose no one would contend that by selling his interest above one hundred feet from the surface he could prevent the State from limiting, by the police power, the height of structures in a city. And why should a sale of underground rights bar the State's power? For aught that appears the value of the coal kept in place by the restriction may be negligible as compared with the value of the whole property, or even as compared with that part of it which is represented by the coal remaining in place and which may be extracted despite the statute. Ordinarily a police regulation, general in operation, will not be held void as to a particular property, although proof is offered that owing to conditions peculiar to it the restriction could not reasonably be applied. But even if the particular facts are to govern, the statute should, in my opinion, be upheld in this case. For the defendant has failed to adduce any evidence from which it appears that to restrict its mining operations was an unreasonable exercise of the police power. Where the surface and the coal belong to the same person, self-interest would ordinarily prevent mining to such an extent as to cause a subsidence. It was, doubtless, for this reason that the legislature, estimating the degrees of danger, deemed statutory restriction unnecessary for the public safety under such conditions.

REVIEW AND DISCUSSION QUESTIONS

1. Justice Holmes writes, "Government could hardly go on if to some extent values incident to property could not be diminished without paying for every such change." Using examples, explain why you agree or disagree.
2. Justice Holmes asserts (1) that the Kohler Act makes the mining of coal "commercially impractical" and (2) that this is the same as appropriating it or destroying it. Is he right on either or both points?
3. In his dissent Justice Brandeis argues (1) that the Kohler Act is "merely the prohibition of a noxious use" and (2) that "its purpose is to protect the public." Is he right on either or both points?
4. Justice Holmes seems to think that the Kohler Act does not represent a strong public interest. In his dissent, Justice Brandeis maintains the opposite. Which of them is right about this? More generally, Holmes believes, and Brandeis denies, that in this case government regulation of property has gone too far. With whom do you agree, and why? Should Pennsylvania have to pay coal companies if it wants to restrict their mining in this way?
5. In your view, in what situations does government regulation of property constitute a "taking" that requires paying the owners compensation? The Fifth Amendment permits the government to take property if "just compensation" is paid. What determines "just compensation"? Are there situations in which it would be wrong of the government to take property even if compensation is paid?

Eminent Domain

Kelo v. City of New London

Eminent domain is the ancient right of government to take property from an individual without consent for the common good—for example, to appropriate someone's land to use for a highway or an airport. The U.S. Constitution recognizes that governmental right, but restricts it. In what is called the "takings clause," the Fifth Amendment states that "private property" shall not "be taken for public use, without just compensation." But what constitutes "public use"? Under what circumstances is the government justified in taking an individual's property for the greater benefit of the community?

---

545 U.S. 469 (2205).

Writing for the majority of the Court, Justice Stevens argues that it was constitutionally permissible for the City of New London to appropriate the homes of Susette Kelo and her neighbors as part of the city's economic development plan. In a robust dissent, Justice O'Connor argues that the majority decision would permit the government to take over any home or business if the property could be put to more advantageous use by someone else. Far from being only a technical legal dispute, *Kelo v. City of London* pits the rights of private property against the good of the community and raises basic questions of economic justice and the limits of government authority.

Justice Stevens delivered the opinion of the Court: In 2000, the city of New London approved a development plan that, in the words of the Supreme Court of Connecticut, was "projected to create in excess of 1,000 jobs, to increase tax and other revenues, and to revitalize an economically distressed city, including its downtown and waterfront areas." In assembling the land needed for this project, the city's development agent has purchased property from willing sellers and proposes to use the power of eminent domain to acquire the remainder of the property from unwilling owners in exchange for just compensation. The question presented is whether the city's proposed disposition of this property qualifies as a "public use" within the meaning of the Takings Clause of the Fifth Amendment to the Constitution.

I

The city of New London (hereinafter City) sits at the junction of the Thames River and the Long Island Sound in southeastern Connecticut. Decades of economic decline led a state agency in 1990 to designate the City a "distressed municipality." In 1996, the Federal Government closed the Naval Undersea Warfare Center, which had been located in the Fort Trumbull area of the City and had employed over 1,500 people. In 1998, the City's unemployment rate was nearly double that of the State, and its population of just under 24,000 residents was at its lowest since 1920.

These conditions prompted state and local officials to target New London, and particularly its Fort Trumbull area, for economic revitalization. To this end, respondent New London Development Corporation (NLDC), a private nonprofit entity established some years earlier to assist the City in planning economic development, was reactivated. In January 1998, the State authorized a $5.35 million bond issue to support the NLDC's planning activities and a $10 million bond issue toward the creation of a Fort Trumbull State Park. In February, the pharmaceutical company Pfizer Inc. announced that it would build a $300 million research facility on a site immediately adjacent to Fort Trumbull; local planners hoped that Pfizer would draw new business to the area, thereby serving as a catalyst to the area's rejuvenation. . . . Upon obtaining state-level approval, the NLDC finalized an integrated development plan focused on 90 acres of the Fort Trumbull area.

The Fort Trumbull area is situated on a peninsula that juts into the Thames River. The area comprises approximately 115 privately owned properties, as well as the 32 acres of land formerly occupied by the naval facility (Trumbull State Park now occupies 18 of those 32 acres). The development plan encompasses seven parcels. Parcel 1 is designated for a waterfront conference hotel at the center of a "small urban village" that will include restaurants and shopping. This parcel will also have marinas for both recreational and commercial uses. A pedestrian "riverwalk" will originate here and continue down the coast, connecting the waterfront areas of the development. Parcel 2 will be the site of approximately 80 new residences organized into an urban neighborhood and linked by public walkway to the remainder of the development, including the state park. This parcel also includes space reserved for a new U.S. Coast Guard Museum. Parcel 3, which is located immediately north of the Pfizer facility, will contain at least 90,000 square feet of research and development office space. Parcel 4A is a 2.4-acre site that will be used either to support the adjacent state park, by providing parking or retail services for visitors, or to support the nearby marina. Parcel 4B will include a renovated marina, as well

as the final stretch of the riverwalk. Parcels 5, 6, and 7 will provide land for office and retail space, parking, and water-dependent commercial uses.

The NLDC intended the development plan to capitalize on the arrival of the Pfizer facility and the new commerce it was expected to attract. In addition to creating jobs, generating tax revenue, and helping to "build momentum for the revitalization of downtown New London," the plan was also designed to make the City more attractive and to create leisure and recreational opportunities on the waterfront and in the park.

The city council approved the plan in January 2000 . . . The city council also authorized the NLDC to purchase property or to acquire property by exercising eminent domain in the City's name. The NLDC successfully negotiated the purchase of most of the real estate in the 90-acre area, but its negotiations with petitioners failed. As a consequence, in November 2000, the NLDC initiated the condemnation proceedings that gave rise to this case.

II

Petitioner Susette Kelo has lived in the Fort Trumbull area since 1997. She has made extensive improvements to her house, which she prizes for its water view. Petitioner Wilhelmina Dery was born in her Fort Trumbull house in 1918 and has lived there her entire life. Her husband Charles (also a petitioner) has lived in the house since they married some 60 years ago. In all, the nine petitioners own 15 properties in Fort Trumbull—4 in parcel 3 of the development plan and 11 in parcel 4A. Ten of the parcels are occupied by the owner or a family member; the other five are held as investment properties. There is no allegation that any of these properties is blighted or otherwise in poor condition; rather, they were condemned only because they happen to be located in the development area. . . .

III

Two polar propositions are perfectly clear. On the one hand, it has long been accepted that the sovereign may not take the property of *A* for the sole purpose of transferring it to another private party *B*, even though *A* is paid just compensation. On the other hand, it is equally clear that a State may transfer property from on private party to another if future "use by the public" is the purpose of the taking; the condemnation of land for a railroad with common-carrier duties is a familiar example. Neither of these propositions, however, determines the disposition of this case.

As for the first proposition, the City would no doubt be forbidden from taking petitioners' land for the purpose of conferring a private benefit on a particular private party. Nor would the City be allowed to take property under the mere pretext of a public purpose, when its actual purpose was to bestow a private benefit. The takings before us, however, would be executed pursuant to a "carefully considered" development plan. The trial judge and all the members of the Supreme Court of Connecticut agreed that there was no evidence of an illegitimate purpose in this case. . . .

On the other hand, this is not a case in which the City is planning to open the condemned land—at least not in its entirety—to use by the general public. Nor will the private lessees of the land in any sense be required to operate like common carriers, making their services available to all comers. But although such a projected use would be sufficient to satisfy the public use requirement, this "Court long ago rejected any literal requirement that condemned property be put into use for the general public." . . . It embraced the broader and more natural interpretation of public use as "public purpose." . . .

The disposition of this case therefore turns on the question whether the City's development plan serves a "public purpose." Without exception, our cases have defined that concept broadly, reflecting our longstanding policy of deference to legislative judgments in this field.

In *Berman* v. *Parker* (1954), this Court upheld a redevelopment plan targeting a blighted area of Washington, D.C., in which most of the housing for the area's 5,000 inhabitants was beyond repair. Under the plan, the area would be condemned and part of it utilized for the construction of streets, schools, and other public facilities. The remainder of the land would be leased or sold to private parties for the purpose of redevelopment, including the construction of low-cost housing.

The owner of a department store located in the area challenged the condemnation, pointing out that his store was not itself blighted and arguing that the creation of a "better balanced, more attractive community" was not a valid public use. Writing for a unanimous Court, Justice Douglas refused to evaluate this claim in isolation, deferring instead to the legislative and agency judgment that the area "must be planned as a whole" for the plan to be successful. The Court explained that "community redevelopment programs need not, by force of the Constitution, be on a piecemeal basis—lot by lot, building by building." The public use underlying the taking was unequivocally affirmed:

> We do not sit to determine whether a particular housing project is or is not desirable. The concept of the public welfare is broad and inclusive. . . . The values it represents are spiritual as well as physical, aesthetic as well as monetary. It is within the power of the legislature to determine that the community should be beautiful as well as healthy, spacious as well as clean, well-balanced as well as carefully patrolled. In the present case, the Congress and its authorized agencies have made determinations that take into account a wide variety of values. It is not for us to reappraise them. If those who govern the District of Columbia decide that the Nation's Capital should be beautiful as well as sanitary, there is nothing in the Fifth Amendment that stands in the way.

In *Hawaii Housing Authority* v. *Midkiff* (1984), the Court considered a Hawaii statute whereby fee title was taken from lessors and transferred to lessees (for just compensation) in order to reduce the concentration of land ownership. We unanimously upheld the statute and rejected the Ninth Circuit's view that it was "a naked attempt on the part of the state of Hawaii to take the property of A and transfer it to B solely for B's private use and benefit." Reaffirming *Berman's* deferential approach to legislative judgments in this field, we concluded that the State's purpose of eliminating the "social and economic evils of a land oligopoly" qualified as a valid public use. Our opinion also rejected the contention that the mere fact that the State immediately transferred the properties to private individuals upon condemnation somehow diminished the public character of the taking. "[I]t is only the taking's purpose, and not its mechanics," we explained, that matters in determining public use. . . .

Viewed as a whole, our jurisprudence has recognized that the needs of society have varied between different parts of the Nation, just as they have evolved over time in response to changed circumstances. Our earliest cases in particular embodied a strong theme of federalism, emphasizing the "great respect" that we owe to state legislatures and state courts in discerning local public needs. For more than a century, our public use jurisprudence has wisely eschewed rigid formulas and intrusive scrutiny in favor of affording legislatures broad latitude in determining what public needs justify the use of the takings power.

IV

Those who govern the City were not confronted with the need to remove blight in the Fort Trumbull area, but their determination that the area was sufficiently distressed to justify a program of economic rejuvenation is entitled to our deference. The City has carefully formulated an economic development plan that it believes will provide appreciable benefits to the community, including—but by no means limited to—new jobs and increased tax revenue. As with other exercises in urban planning and development, the City is endeavoring to coordinate a variety of commercial, residential, and recreational uses of land, with the hope that they will form a whole greater than the sum of its parts. To effectuate this plan, the City has invoked a state statute that specifically authorizes the use of eminent domain to promote economic development. . . . Because that plan unquestionably serves a public purpose, the takings challenged here satisfy the public use requirement of the Fifth Amendment.

To avoid this result, petitioners urge us to adopt a new bright-line rule that economic development does not qualify as a public use. . . . Neither precedent nor logic supports petitioners' proposal. Promoting economic development is a traditional and long accepted function of government. There is, moreover, no principled way of distinguishing economic development from the other public purposes that we have recognized. . . .

It is further argued that without a bright-line rule nothing would stop a city from transferring citizen *A*'s property to citizen *B* for the sole reason that citizen *B* will put the property to a more productive use and thus pay more taxes. Such a one-to-one transfer of property, executed outside the confines of an integrated development plan, is not presented in this case. While such an unusual exercise of government power would certainly raise a suspicion that a private purpose was afoot, the hypothetical cases posited by petitioners can be confronted if and when they arise. They do not warrant the crafting of an artificial restriction on the concept of public use. . . .

In affirming the City's authority to take petitioners' properties, we do not minimize the hardship that condemnations may entail, notwithstanding the payment of just compensation. We emphasize that nothing in our opinion precludes any State from placing further restrictions on its exercise of the takings power. Indeed, many States already impose "public use" requirements that are stricter than the federal baseline. Some of these requirements have been established as a matter of state constitutional law, while other are expressed in state eminent domain statutes that carefully limit the grounds upon which takings may be exercised. . . . The necessity and wisdom of using eminent domain to promote economic development are certainly matters of legitimate public debate. This Court's authority, however, extends only to determining whether the City's proposed condemnations are for a "public use" within the meaning of the Fifth Amendment to the Federal Constitution. Because over a century of our case law interpreting that provision dictates an affirmative answer to that question, we may not grant petitioners the relief that they seek.

The judgment of the Supreme Court of Connecticut is affirmed.

It is so ordered.

Justice O'Connor, with whom the Chief Justice, Justice Scalia and Justice Thomas join, dissenting: Over two centuries ago, just after the Bill of Rights was ratified, Justice Chase wrote:

> An *act* of the Legislature (for I cannot call it a law) contrary to the great first principles of the social compact, cannot be considered a rightful exercise of legislative authority. . . . A few instances will suffice to explain what I mean. . . . [A] law that takes property from A. and gives it to B: It is against all reason and justice, for a people to entrust a Legislature with *such* powers; and therefore, it cannot be presumed that they have done it.

Today the Court abandons this long-held, basic limitation on government power. Under the banner of economic development, all private property is now vulnerable to being taken and transferred to another private owner, so long as it might be upgraded—*i.e.*, given to an owner who will use it in a way that the legislature deems more beneficial to the public—in the process. To reason, as the Court does, that the incidental public benefits resulting from the subsequent ordinary use of private property render economic development takings "for public use" is to wash out any distinction between private and public use of property—and thereby effectively to delete the words "for public use" from the Takings Clause of the Fifth Amendment. Accordingly I respectfully dissent. . . .

I

The Fifth Amendment to the Constitution, made applicable to the States by the Fourteenth Amendment, provides that "private property [shall not] be taken for public use, without just compensation." When interpreting the Constitution, we begin with the unremarkable presumption that every word in the document has independent meaning, "that no word was unnecessarily used, or needlessly added." In keeping with that presumption, we have read the Fifth Amendment's language to impose two distinct conditions on the exercise of eminent domain: "the taking must be for a 'public use' and 'just compensation' must be paid to the owner."

These two limitations serve to protect "the security of Property," which Alexander Hamilton described to the Philadelphia Convention as one of the "great obj[ects] of Gov[ernment]." Together they ensure stable property ownership by providing safeguards against excessive, unpredictable, or unfair use of the government's eminent domain power—particularly against those

owners who, for whatever reasons, may be unable to protect themselves in the political process against the majority's will.

While the Takings Clause presupposes that government can take private property without the owner's consent, the just compensation requirement spreads the cost of condemnations and thus "prevents the public from loading upon one individual more than his just share of the burdens of government." The public use requirement, in turn, imposes a more basic limitation, circumscribing the very scope of the eminent domain power: Government may compel an individual to forfeit her property for the *public's* use, but not for the benefit of another private person. This requirement promotes fairness as well as security.

Where is the line between "public" and "private" property use? We give considerable deference to legislatures' determinations about what governmental activities will advantage the public. But were the political branches the sole arbiters of the public-private distinction, the Public Use Clause would amount to little more than hortatory fluff. An external, judicial check on how the public use requirement is interpreted, however limited, is necessary if this constraint on government power is to retain any meaning.

Our cases have generally identified three categories of takings that comply with the public use requirement, though it is in the nature of things that the boundaries between these categories are not always firm. Two are relatively straightforward and uncontroversial. First, the sovereign may transfer private property to public ownership—such as for a road, a hospital, or a military base. Second, the sovereign may transfer private property to private parties, often common carriers, who make the property available for the public's use—such as with a railroad, a public utility, or a stadium. But "public ownership" and "use-by-the-public" are sometimes too constricting and impractical ways to define the scope of the Public Use Clause. Thus we have allowed that, in certain circumstances and to meet certain exigencies, takings that serve a public purpose also satisfy the Constitution even if the property is destined for subsequent private use.

This case returns us for the first time in over 20 years to the hard question of when a purportedly "public purpose" taking meets the public use requirement. It presents an issue of first impression: Are economic development takings constitutional? I would hold that they are not. We are guided by two precedents about the taking of real property by eminent domain. In *Berman*, we upheld takings within a blighted neighborhood of Washington, D.C. The neighborhood had so deteriorated that, for example, 64.3% of its dwellings were beyond repair. . . .

In *Midkiff*, we upheld a land condemnation scheme in Hawaii whereby title in real properly was taken from lessors and transferred to lessees. At that time, the State and Federal Governments owned nearly 49% of the State's land, and another 47% was in the hands of only 72 private landowners. Concentration of land ownership was so dramatic that on the State's most urbanized island, Oahu, 22 landowners owned 72.5% of the fee simple titles. The Hawaii Legislature had concluded that the oligopoly in land ownership was "skewing the State's residential fee simple market, inflating land prices, and injuring the public tranquility and welfare," and therefore enacted a condemnation scheme for redistributing title. . . .

The Court's holdings in *Berman* and *Midkiff* were true to the principle underlying the Public Use Clause. In both those cases, the extraordinary, precondemnation use of the targeted property inflicted affirmative harm on society—in *Berman* through blight resulting from extreme poverty and in *Midkiff* through oligopoly resulting from extreme wealth. And in both cases, the relevant legislative body had found that eliminating the existing property use was necessary to remedy the harm. Thus a public purpose was realized when the harmful use was eliminated. Because each taking *directly* achieved a public benefit, it did not matter that the property was turned over to private use. Here, in contrast, New London does not claim that Susette Kelo's and Wilhelmina Dery's well-maintained homes are the source of any social harm. Indeed, it could not so claim without adopting the absurd argument that any single-family home that might be razed to make way for an apartment building, or any church that might be replaced with a retail store, or any small business that might be more lucrative if it were instead part of a national franchise, is inherently harmful to society and thus within the government's power to condemn.

In moving away from our decisions sanctioning the condemnation of harmful property use, the Court today significantly expands the meaning of public use. It holds that the sovereign may take private property currently put to ordinary private use, and give it over for new, ordinary private use, so long as the new use is predicted to generate some secondary benefit for the public—such as increased tax revenue, more jobs, maybe even aesthetic pleasure. But nearly any lawful use of real private property can be said to generate some incidental benefit to the public. Thus, if predicted (or even guaranteed) positive side-effects are enough to render transfer from one private party to another constitutional, then the words "for public use" do not realistically exclude *any* takings, and thus do not exert any constraint on the eminent domain power. . . . Who among us can say she already makes the most productive or attractive possible use of her property? The specter of condemnation hangs over all property. Nothing is to prevent the State from replacing any Motel 6 with a Ritz-Carlton, any home with a shopping mall, or any farm with a factory.

The Court also puts special emphasis on facts peculiar to this case: The NLDC's plan is the product of a relatively careful deliberative process; it proposes to use eminent domain for a multipart, integrated plan rather than for isolated property transfer; it promises an array of incidental benefits (even aesthetic ones), not just increased tax revenue; it comes on the heels of a legislative determination that New London is a depressed municipality. But none has legal significance to blunt the force of today's holding. If legislative prognostications about the secondary public benefits of a new use can legitimate a taking, there is nothing in the Court's rule . . . to prohibit property transfers generated with less care, that are less comprehensive, that happen to result from less elaborate process, whose only projected advantage is the incidence of higher taxes, or that hope to transform an already prosperous city into an even more prosperous one. Finally, in a coda, the Court suggests that property owners should turn to the States, who may or may not choose to impose appropriate limits on economic development takings. This is an abdication of our responsibility. States play many important functions in our system of dual sovereignty, but compensating for our refusal to enforce properly the Federal Constitution (and a provision meant to curtail state action, no less) is not among them.

***

It was possible after *Berman* and *Midkiff* to imagine unconstitutional transfers from A to B. Those decisions endorsed government intervention when private property use had veered to such an extreme that the public was suffering as a consequence. Today nearly all real property is susceptible to condemnation on the Court's theory. . . . Any property may now be taken for the benefit of another private party, but the fallout from this decision will not be random. The beneficiaries are likely to be those citizens with disproportionate influence and power in the political process, including large corporations and development firms. As for the victims, the government now has license to transfer property from those with fewer resources to those with more. The founders cannot have intended this perverse result. "[T]hat alone is a *just* government," wrote James Madison, "which *impartially* secures to every man, whatever is his *own*."

REVIEW AND DISCUSSION QUESTIONS

1. Describe the facts of this case. What are the main features of New London's economic development plan? What was its motivation in adopting it?
2. As described by Justice Stevens, what are the "two polar propositions" that provide the legal background to this case?
3. Both Justice Stevens and Justice O'Connor discuss the Court's earlier *Berman* and *Midkiff* decisions. What was decided in those cases? Those cases seem to support the majority's position in *Kelo*. Justice O'Connor, however, believes that the situation in *Kelo* differs significantly from them. Who is right about this? Is O'Connor's dissent consistent with those earlier decisions?

4. How does Justice Stevens respond to the worry that a city might transfer "citizen A's property to citizen B for the sole reason that citizen B will put the property to a more productive use"?
5. Justice O'Connor gives two "straightforward and uncontroversial" examples of government takings that comply with the public use requirement. What are they?
6. At the end of his decision, Justice Stevens points out that the majority's decision does not prevent individual states from imposing stricter "public use" requirements or limiting in other ways the right of communities to take private property. How does Justice O'Connor respond to this?
7. Justices Stevens and O'Connor disagree whether New London's appropriation of Susette Kelo's home was a proper exercise of government power. With whom do you agree and why?
8. Did New London treat Susette Kelo and her neighbors fairly? Assuming that the proposed development will help to revitalize New London, is it reasonable for the city to appropriate private property to carry out the plan? Is doing so just? Was it reasonable or fair of Kelo to refuse to sell and thus obstruct the city's development plan?
9. "If 'just compensation' is paid, then by definition those who lose their property cannot claim that they have been treated unjustly." Assess this argument. Can compensation be just if one of the parties is unwilling to accept it?

CHAPTER 18

Private Agreements: The Law of Contract

In any socioeconomic system based on private property and market exchange, agreements between private parties will play an extensive and important role. It is not surprising, then, that contract law is a main branch of the civil law. Whereas tort law concerns violations of various legally recognized duties, in contracts the parties themselves, in large measure, determine through mutual agreement what their obligations toward one another will be. By contrast with the criminal law, contractual obligations are not only self-imposed but also privately enforced: litigation begins only when one party to the contract believes that the other has failed to live up to the terms of the agreement.

As with property law, philosophical disagreements over the moral basis of contract, and the nature of the obligation that contracts create, pervade legal debates over the adjudication of particular cases and the interpretation of contract law in general. These disagreements are reflected in the readings that follow. Morris Cohen explains the historical context of contract law and emphasizes its institutional aspects. Although contracts are private agreements, they have broader ramifications and grow into social and legal institutions. Society, therefore, has a legitimate interest in imposing certain standards and shaping contracts in ways that meet its goals. Charles Fried stands in an opposed tradition. He sees contractual obligation as rooted in promise. In his view, it is the moral force of promising—not social utility—that is the basis of contract law and that should guide the way the courts adjudicate contract disputes.

This contrast of perspectives plays out in the cases that follow. The *Walker-Thomas Furniture* case raises two questions: what it means for a contract to be "unconscionable" and whether courts should refuse to enforce the terms of such contracts. The *Baby M* case concerns the legality of surrogacy contracts. Although no one doubts that both parties to this case intentionally entered into a formal compact, the question remains whether the law should recognize and uphold their consensual agreement as a binding contract. The famous *Lochner* case, and the cases that accompany it, deal with the right of the government to impose limits and restrictions on employment contracts into which private parties have entered.

The Basis of Contract

Morris Raphael Cohen

Contract law is complicated, and its details concern the conditions that must be met for a contract to be valid and legally enforceable as well as, of course, the various remedies available to those damaged by a breach of contract. In this essay, however, Morris Raphael Cohen discusses the nature of contract as a legal and social institution from a broader historical and philosophical perspective. After examining the social and historical roots of contract and assessing the political and economic arguments for what he calls contractualism, Cohen emphasizes the extent to which contract law can be seen as a branch of public law. Because the law must have regard for the general and institutional effects of classes of transactions, it must, when settling disputes, be prepared to go beyond the original intention of the contracting parties.

The nature of contract has been much discussed by lawyers interested in specific technical doctrines, and by moralists, economists, and political theorists interested in general social philosophy. There is still need for some effort to combine these points of view. The bearings of general philosophy become more definite through its applications, and the meaning of a technical doctrine receives illumination when we see it in the light of those wider ideas of which it is the logical outcome.

This large and important task is obviously beyond the limits of a short paper. But a few suggestions may indicate something of the scope of the problem.

1. THE SOCIAL ROOTS OF CONTRACT

One of the most influential of modern saws is Maine's famous dictum that the progress of the law has been from status to contract. It has generally been understood as stating not only a historical generalization but also a judgment of sound policy—that a legal system wherein rights and duties are determined by the agreement of the parties is preferable to a system wherein they are determined by "status."

. . . That Maine's generalization is not a universal and necessary law, he himself recognized in his treatment of feudal land tenure. The rights and duties of sovereign and subject, of homage or fealty and protecting lordship, were contractual in the early Middle Ages, and gradually ceased to be so as they became customary and were later replaced by the legislation of the modern national states. It is also true . . . that the modern state has, in all civilized countries, been steadily increasing the scope of its functions, so that men now do things by virtue of their status as citizens and taxpayers which formerly they did by voluntary agreement. One only needs to mention the fields of charity and education to make this obvious. Moreover, in many relations in which men are more or less free to enter into contracts, such as that between insurer and insured, landlord and tenant, employer and employee, shipper and carrier, the terms of the agreement are more and more being fixed by law, so that the entering into these relations has something analogous to the entering into the relation of marriage, trusteeship, or public office. The specific rights and duties are not fixed by agreement, though the assumption of the relations is more or less voluntary.

Nevertheless there is enough truth in Maine's observation to warrant a more discriminating attitude to it than that of complete acceptance or complete rejection.

Looking at the matter macroscopically rather than microscopically, there can be little doubt that legally binding agreements or promises play a smaller part in the earlier history of all known peoples. The development of contract is largely an incident of commercial and industrial enterprises that involve a greater anticipation of the future than is necessary in a simpler or more primitive economy. In the latter the solidarity of

relatively self-sufficient family groups and the fear of departing from accustomed ways limit individual initiative as well as the scope and importance of what can be achieved by deliberate agreements or bargains. In some respects, however, less developed societies resort more than we do to contracts or compacts and enforce promises that we no longer enforce. Thus they preserve peace not by organized police or standing armies, but by agreements like our present treaties of peace between nations; and promises to the gods, which are now matters of individual conscience, used to be enforced by the community as a whole because it feared the undiscriminating effects of divine wrath. . . .

2. THE POLITICAL THEORY OF CONTRACTUALISM

Contractualism in the law, that is, the view that in an ideally desirable system of law all obligation would arise only out of the will of the individual contracting freely, rests not only on the will theory of contract but also on the political doctrine that all restraint is evil and that the government is best which governs least. This in turn is connected with the classical economic optimism that there is a sort of preestablished harmony between the good of all and the pursuit by each of his own selfish economic gain. . . .

The argument that a régime of free contract assures the greatest amount of liberty for all is characteristic of the eighteenth century philosophy of the Enlightenment and is still essential to the faith of Jeffersonian democracy behind our bills of rights. The older Calvinistic argument for government rested on the need of restraining the wickedness of man (due to the corruption of the flesh) by rules and magistrates deriving their power from God. Against this the deistic and bourgeois Enlightenment developed the contrary view, that men are inherently good and that their dark deeds have been due to the corruption and superstition brought about by tyrants and priests. As we get rid of the latter, the original, benevolent nature of man asserts itself and history indeed shows a gradual but steady progress in the direction of freedom. It was natural for the representatives of the growing commercial and industrial interests to view the state, controlled as it had been by landed barons and prelates (lords temporal and spiritual), as exclusively an instrument of oppression, and necessarily evil. But their argument overshot its mark. They forgot that not only industry but also the whole life of civilization depends on the feeling of security that the protection of the government or organized community affords. . . .

Since what we want generally seems to us good, freedom, as the removal of obstacles to achievement, is a necessary part or condition of this good. But mere freedom as absence of restraint, without positive power to achieve what we deem good, is empty and of no real value. The freedom to make a million dollars is not worth a cent to one who is out of work. Nor is the freedom to starve, or to work for wages less than the minimum of subsistence, one that any rational being can prize—whatever learned courts may say to the contrary. . . .

In the United States, the Jeffersonian democrats fought against the power of the government both in the nation and in their states. They feared that the merchants and the large landowners of the seaboard, who had controlled the older states, would control the central government. To limit the power of government by a system of checks and balances was therefore the way to assure liberty from oppression. So long as the country was sparsely settled and our people remained for the most part a nation of independent freeholders, this was a workable theory. It never, however, was carried out consistently. For the temptation to use the government for positive ends, such as education or the safeguarding of our commercial interests in the West and on the seas, could not be resisted even by Jefferson himself. In the last forty or fifty years the representatives of large industry have invoked this theory of the bills of rights to limit legislative power to regulate industry in the interests not only of the workers but also of the future manhood and womanhood of the nation. The same group, however, that protests against a child labor law, or against any minimum wage law intended to insure a minimum standard of decent living, is constantly urging the government to protect industry by tariffs. Clearly, the theory of *laissez faire*, of complete non-interference of the government in business, is not really held consistently by those who so frequently invoke it. A government so limited in its

powers that it could do no harm would be useless, since it could do no good. . . .

The Economic Argument for Contractualism

When the political argument is closely pressed, it is found to rest on the economic one that a regime in which contracts are freely made and generally enforced gives greater scope to individual initiative and thus promotes the greatest wealth of a nation. . . .

The clearest and most convincing statement of the case for the classical theory of free competition is that of Mr. Justice Holmes. Let us, he urges, get behind the fact of ownership, and look at the processes of production and consumption of goods. The men who achieve great private fortunes do not consume very much of this social wealth. Their fortunes denote rather power to control the flow of goods. And who is better fitted to command this process of production and distribution than the man who wins it in the competition of the market? The assumption behind this is that the man who succeeds in winning a fortune (not, it should be noted, the man who receives it by inheritance) has succeeded because he has been able to anticipate the largest effective demand for goods and to organize the most economical way of producing them.

One weakness of this argument is that it ignores the frightful waste involved in competition. The community as a whole ultimately pays the cost, in labor and capital goods (including their extensive sales and advertising forces), of all the economic enterprises that are allowed to compete and fail. Moreover, the greatest profits do not always come with the greatest productivity. There are monopoly profits, like the unearned increment of land value, that clearly do not arise from productivity of the owners, and there are monopoly profits that are swelled by reducing the output, so that fishermen, wheat and cotton growers, and other producers are often advised to do this. Neither can free competition prevent the paradoxical situation that our economic crises repeatedly show, namely, an overstocked food market and general destitution from inability to buy. The latter is certainly in part due to the fact that, under unrestrained competition, wages and the return for the labor of the farmer are not sufficient to enable the vast majority of the people to buy enough of what they have produced. Thus, some of the supposedly greater efficiency of private over public business, to the extent that it involves lower real wages, is detrimental to the general welfare. The latter depends not only on the mass of production, but also on the kind of goods produced, on the conditions under which men work, and on the ways in which the product is distributed.

For these reasons it is rare nowadays to find any advocates of a régime of free competition except among certain lawyers and judges who use it to oppose regulation of the "labor contract" by the state. The general consensus among business men has demanded the organization of our Interstate Commerce Commission and Federal Trade Commission, our state railway and public service commissions, our state insurance and industrial commissions, and other administrative bodies that limit and regulate certain essential business contracts. Also the great captains of industry are everywhere trying to eliminate free competition. And those who talk about "keeping the government out of business" are the last to desire that the government shall not help or protect, by proper rules, the business in which they are involved. The differences that divide men in this respect concern the questions of what interests should be protected and who should control the government. . . .

Excesses of Contractualism

As the result of the various forces that have thus supported the cult of contractualism there has been developed in all modern European countries (and in those which derive from them) a tendency to include within the categories of contract transactions in which there is no negotiation, bargain, or genuinely voluntary agreement. Let us consider a few typical situations.

A citizen going to work boards a street car and drops a coin in the conductor's or motorman's box. This, or the buying of a ticket, is treated as a contract, and courts and jurists speak of its "terms" and of the rights and duties under it. No one claims that there is any actual "meeting of minds" of the passenger and the street railway corporation.

There is no actual offer or acceptance—certainly no bargaining between the two parties. The rights and duties of both are prescribed by law and are the same no matter what, if anything, goes on in the passenger's mind, or in the corporation's, if it has any mind. Moreover, the liabilities of the railway corporation to the passenger are in many circumstances exactly the same even if he does not buy any ticket. Obviously, therefore, we have here a situation in which the law regulates the relation between different parties and it is pure fiction to speak of it as growing out of any agreement of the wills of the parties.

A more serious confusion of fact and fiction occurs when we speak of the "labor contract." There is, in fact, no real bargaining between the modern large employer (say the United States Steel Corporation) and its individual employees. The workingman has no real power to negotiate or confer with the corporation as to the terms under which he will agree to work. He either decides to work under the conditions and schedule of wages fixed by the employer or else he is out of a job. If he is asked to sign any paper he does so generally without any knowledge of what it contains and without any real freedom to refuse. For we cannot freely change our crafts, and if a man is a weaver or shoe laster, he is dependent on the local carpet or shoe factory for his livelihood, especially so if he has a family, which is not as mobile as money. The greater economic power of the employer exercises a compulsion as real *in fact* as any now recognized by law as duress. The extreme form of such duress, the highwayman's pistol, still leaves us with the freedom to accept the terms offered or else take the consequences. But such choice is surely the very opposite of what men value as freedom.

Clearly, then, the element of consent on the part of the employee may be a minor one in the relation of employment—a relation much more aptly and realistically described by the old law as that between master and servant. Down to the end of the eighteenth century this relation was in fact regulated by the government. Wages used to be regularly fixed by justices of the peace under the authority of parliamentary enactments, and even the beer that the master was to serve to the servant with his bread had its strength regulated by law. Any demand by workmen for higher wages or any accession to such demands on the part of masters was a violation of the law. Yet courts now speak as if the effort on the part of the state to regulate wages were an unheard of interference with the eternal laws of nature. . . . [But] only after the Civil War [was] the doctrine [invented] that the "right *to* contract" [is a constitutionally protected property right]. . . .

The spread of contractualistic notions shows itself in the tendency to speak of marriage as itself a contract. Now there are, usually, solemn promises exchanged when the marriage ceremony is performed and there may be agreements as to dowry and other property rights. But the specific legal relations of husband and wife are by no means determined thereby. These relations are entirely fixed by law and the parties to it cannot vary its terms, just as they cannot vary the terms of their obligations to any children they may bring into the world. If there is no sense in speaking of the rights and duties between parents and minor children as contractual, neither is there in speaking of the relations of husband and wife as contractual. The fact that an act is more or less voluntary does not make its legal consequences contractual.

The extreme of contractualistic thought was reached when European publicists of the eighteenth century and American judges of the nineteenth century spoke of the social compact or contract as the basis of society and of all law.

I do not want to add to the many (in fact, too many) refutations of the social-contract theory. . . . The tradition of Hebrew history and Greek philosophy that bases government and law upon covenant or agreement has had a salutary influence in challenging all law to justify itself at the bar of reason. It is good to ask of any law whether it is such as rational beings would adopt if they wanted to establish a society. Nevertheless, there are obviously insuperable difficulties in trying to derive all legal obligations from contract. Children have certain obligations to their parents that are not contractual. Indeed, we may well ask, "Why should we obey laws that our ancestors and not we agreed to?" We may even go further and ask, "Why *should* we keep agreements that we made some time ago when we were younger and less experienced or wise than we are now?" If there is any rational answer to either of these queries, it must take the

form of indicating some social good or necessity that is served by our keeping our promises. But if so, why may not the same social good or necessity be served by making children obey and at times support their parents, or by making those who hold property pay for sewage, education, and other communal necessities, even if they do not agree? The merits of the issue are not really affected by introducing a fictional contract.

These attempts to stretch the category of contract err in failing to recognize a certain necessary social solidarity, especially that of any generation with its ancestors, as not the outcome but the very basis of contract itself. . . .

3. CONTRACT AND SOVEREIGNTY

. . . The cardinal error of the traditional individualistic theories of contract is their way of speaking as if the law does nothing but put into effect what the contracting parties originally agreed on. The best that can be said for this is that it may sometimes be true. But even if that were more generally the case, we should still have to attach more importance to the factor of enforcement than the prevailing theories do. The fact that two people agree to do something not prohibited by the public criminal law and carry out their agreement, or fail to do so, does not of itself bring the law of contract into being. A large number of important agreements, even in business, as in social, political, and religious matters, are left to be directly regulated by other agencies, such as the prevailing sense of honor, individual conscience, or the like. It is an error then to speak of the law of contract as if it merely allows people to do things. The absence of criminal prohibition will do that much. The law of contract plays a more positive role in social life, and this is seen when the organized force of the state is brought into play to compel the loser of a suit to pay or to do something. Doubtless most people live up to their promises or agreements either through force of custom or because it is in the long run more advantageous to do so. But there can be no doubt that the possibility of the law's being invoked against us if we fail to do so is an actual factor in the situation. Even if the transactions that come to be litigated are atypical, their judicial determination is still influential in molding the legal custom. For the ruling in a case that departs from the mode supports or opposes some direction of variation and thus fixes the direction of growth of what becomes customary. The fact, then, that in the general run of transactions people do not resort to actual litigation, is certainly in part due to the fact that they know in a general way what will be the outcome of that process. . . .

From this point of view the law of contract may be viewed as a subsidiary branch of public law, as a body of rules according to which the sovereign power of the state will be exercised as between the parties to a more or less voluntary transaction.

The first rules of public law, generally called constitutional law, regulate the conduct of the chief state officials by indicating the scope of their powers. Within this scope legislatures use their discretion or wisdom to enact certain statutes; and judges, by following precedents, elaborate certain rules as to when and how the power of the state shall be exercised. Among these rules we have the laws of partnership, leases, agreements for services, contracts of surety or insurances, and the like. Now, just as the rules of constitutional law are general and leave blanks to be filled in by the legislature, courts, and administrative officials (whose rules and habitual practices are law to those over whom they have authority), so do the rules of contracts allow men to formulate for themselves, within the prescribed limits, certain rights and duties governing certain transactions between them; and when the parties have thus formulated their agreements, the latter become a part of the law of the land, just as much as do treaties between our nation and others, compacts between states, contracts between a state or division thereof and a private corporation, or the grant of a pension to the widow of a former president. When a state or a municipality makes a contract with a public service corporation for gas or transportation at a given price to the consumer, no one doubts that such an agreement is part of the legal order. But so are private agreements that the law sanctions. Thus, when a trade union makes an agreement with an association of employers, or even with a single employer, the result is law not only for those "represented" at the signing of the papers but for all those who wish to enter the industry at any time that the agreement is in force. This is in general true of all more or less permanently organized partnerships, companies, corporations, or other groups; and enforceable

agreements between individuals, no matter on how limited a scale, are similarly part of the law by virtue of the general rules of state action that apply to them.

If, then, the law of contract confers sovereignty on one party over another (by putting the state's forces at the disposal of the former), the question naturally arises: For what purposes and under what circumstances shall that power be conferred? Adherents of the classical theory have recognized that legal enforcement serves to protect and encourage transactions that require credit or reliance on the promises of others. But we also need care that the power of the state be not used for unconscionable purposes, such as helping those who exploit the dire need or weaknesses of their fellows. Usury laws have recognized that he who is under economic necessity is not really free. To put no restrictions on the freedom to contract would logically lead not to a maximum of individual liberty but to contracts of slavery, into which, experience shows, men will "voluntarily" enter under economic pressure—a pressure that is largely conditioned by the laws of property. Regulations, therefore, involving some restrictions on the freedom to contract are as necessary to real liberty as traffic restrictions are necessary to assure real freedom in the general use of our highways.

From this point of view, the movement to standardize the forms of contract—even to the extent of prohibiting variations or the right to "contract out"—is not to be viewed as a reaction to, but rather as the logical outcome of, a régime of real liberty of contract. It is a utilization of the lessons of experience to strengthen those forms which best serve as channels through which the life of the community can flow most freely. . . .

The function of the law of contract in promoting the standardization of transactions is at all times an important one. And the more developed and complicated transactions become, the more there is need for eliminating as much uncertainty as possible by standardization. This is certainly true today. Consider the case of a man who wants to publish a book, to buy an insurance policy or a letter of credit, to ship his goods or to store them in a warehouse, to lease an apartment, to have gas or electricity or telephone service supplied to him, to mortgage his house, or to obtain a surety bond—in all these and in many other relations his freedom to contract is facilitated by standard forms molded by past law and custom. Naturally, standardized contracts, like other laws, serve the interests of some better than those of others; and the question of justice thus raised demands the attention not only of legislatures but also of courts that have to interpret these standard forms and of administrative bodies that have to supervise their enforcement. In a changing social order these standards or forms must grow or become modified; and to make them function more serviceably it is not sufficient to wait until trouble develops and is brought before the courts for adjudication. The need of intelligent anticipation that can be effected by initiating inquiries cannot be met by our traditional court procedure, and this has compelled the joining of administrative with judicial power in the hands of bodies like the Interstate Commerce Commission and our various state public service commissions.

A realization of the growth of standard forms suggests the introduction of a point of view in the study of contract similar to what has been called the institutional approach in the study of economics. . . . A similar change of approach in the study of the law of contract means beginning not with the bargaining between the two parties, but with the legal form or way of doing things, with the established institution within which negotiation is possible.

. . . Contracts are voluntary, fixed, and temporary, while institutions are socially hereditary, grow, and last longer. Yet . . . contracts, especially collective ones, grow into institutions. The marriage relation shows the passage from one to the other. . . . [We should reject] the absolute separation of freedom of contract from government regulation, the former conceived as purely negative and the latter as purely arbitrary. In actual life real freedom to do anything, in art as in politics, depends upon acceptance of the rules of our enterprise. As has been remarked elsewhere, the rules of the sonnet do not hamper real poets but rather help weak ones. Real or positive freedom depends upon opportunities supplied by institutions that involve legal regulation. Our legislative forces may be narrowly partisan and the rules may be poor ones. But this can be remedied not by the abrogation of all rules but by the institution of better ones.

For this reason the notion that government rests on contract—a notion that runs through both our

Hebrew and Greek heritage, and largely conforms to our peculiar American experience—contains a partial truth that should not be utterly disregarded because of some poor arguments in its behalf. If we discard the notion that *all* organized society began in a voluntary contract—a proposition that few have advanced as a literal truth—we may yet recognize that as men become more enlightened they can treat government as if it were a contractual affair, that is, judge the services of governmental rules by the price we pay for them. The great men who founded the rationalistic legal and political tradition of the Enlightenment, Althusius, Grotius, Leibniz, and Locke, may have underestimated the force of tradition, but in treating governmental rights and duties under the categories of contract, they helped to liberalize and humanize our international and our criminal law, as well as the law of private and commercial transactions.

There is no inherent reason for rejecting the view that the roots of the law of contract are many rather than one. Agreements and promises are enforced to enable people to rely on them as a rule and thus make the path of enterprise more secure; but in this connection the law must also go beyond the original intention of the parties to settle controversies as to the distribution of gains and losses that the parties did not anticipate in the same way. Some recognition must always be given to the will or intention of those who made the contract, but the law must always have regard for the *general* effects of classes of transactions, and it cannot free men from the necessity of acting at their peril when they do not know the consequences that the law will attach to their acts—and this needs to be emphasized in any attempt to formulate a rational theory. The law is a going concern and like all social institutions is governed by habit. It therefore will continue to enforce promises and agreements, for no better reason than that they have been enforced and there is no sufficient countervailing consideration to force or justify a break with the established habit that has become the basis of social expectancies. Legal and other habits are not always deliberately formed to serve a definite purpose. Certain forms or ceremonies arise under special circumstances but continue to appeal to us through the principle of economy of effort: it is generally easier to use the existing forms than to break with them and adopt new ones. Of course old forms may become inconvenient or positive hindrances. They are then whittled away by pious fiction or violently changed by revolutionary legislation. In general, however, the ancient truth that men are creatures of habit will put us on guard against the vain assumption that we can get rid of formalism in the law of contract or anywhere else. We may flatter ourselves on getting rid of seal or other ancient binding ceremony. But we must remember that these forms seemed as naturally obligatory to our fathers as the signing of papers or the administering of oaths seem to us today.

In arguing for their indispensability we may recognize that not all forms are perfectly congenial or responsive to the need of the life that pulses through them. And as men become more enlightened they become more ready to discard, as well as employ, diverse instruments or vessels. Wisdom is not attained either by blind acceptance or blind rejection. We need a discriminating evaluation of what exists and what is possible.

REVIEW AND DISCUSSION QUESTIONS

1. Cohen refers to Maine's thesis that there has been an evolution in the law from status to contract. What does this mean, and to what extent is it true?
2. What is the political theory of contractualism? What is the economic argument for contractualism? What reservations does Cohen have about both of them? More generally, what does he see as the "excesses" of contractualism? Can more be said on behalf of contractualism than Cohen says?
3. What does Cohen mean when he writes that "the law of contract may be viewed as a subsidiary branch of public law"? What are the implications of this? What would an "institutional" approach to the study of contract involve?
4. What is important about the standardization of contracts? How can standardization and, more generally, the acceptance of rules promote freedom?

Contract as Promise

Charles Fried

Legal theorists like Morris Cohen emphasize the social and institutional aspects of contract law. As they see it, contracts are private agreements that have broad ramifications and evolve into social institutions. Society, therefore, needs to try to impose standards and shape contracts to meet its goals. The conclusion these writers draw is that the law, rather than simply and neutrally upholding private agreements freely arrived at by the contracting parties, imposes on them legal obligations based on society's own needs.

In this selection from his book *Contract as Promise*, Charles Fried, professor of law at Harvard University and former U.S. solicitor general, upholds a different view of contract, which sees contractual obligation as based on the moral force of promise rather than social or institutional utility. He analyzes the nature of promise and promissory obligation and defends his view of contract as promise against various objections.

THE MORAL OBLIGATION OF PROMISE

Once I have invoked the institution of promising, why exactly is it wrong for me then to break my promise?....

The obligation to keep a promise is grounded not in arguments of utility but in respect for individual autonomy and in trust.... An individual is morally bound to keep his promises because he has intentionally invoked a convention whose function it is to give grounds—moral grounds—for another to expect the promised performance. To renege is to abuse a confidence he was free to invite or not, and which he intentionally did invite. To abuse that confidence now is like (but only *like*) lying: the abuse of a shared social institution that is intended to invoke the bonds of trust. A liar and a promise-breaker each *use* another person. In both speech and promising there is an invitation to the other to trust, to make himself vulnerable; the liar and the promise-breaker then abuse that trust. The obligation to keep a promise is thus similar to but more constraining than the obligation to tell the truth. To avoid lying you need only believe in the truth of what you say when you say it, but a promise binds into the future, well past the moment when the promise is made. There will, of course, be great social utility to a general regime of trust and confidence in promise and truthfulness. But this just shows that a regime of mutual respect allows men and women to accomplish what in a jungle of unrestrained self-interest could not be accomplished. If this advantage is to be firmly established, there must exist a ground for mutual confidence deeper than and independent of the social utility it permits.

The utilitarian counting the advantages affirms the general importance of enforcing *contracts*. The moralist of duty, however, sees *promising* as a device that free, moral individuals have fashioned on the premise of mutual trust, and which gathers its moral force from that premise. The moralist of duty thus posits a general obligation to keep promises, of which the obligation of contract will be only a special case—that special case in which certain promises have attained legal as well as moral force. But since a contract is first of all a promise, the contract must be kept because a promise must be kept.

To summarize: There exists a convention that defines the practice of promising and its entailments. This convention provides a way that a person may create expectations in others. By virtue of the basic Kantian principles of trust and respect, it is wrong to invoke that convention in order to make a promise, and then to break it.

WHAT A PROMISE IS WORTH

If I make a promise to you, I should do as I promise; and if I fail to keep my promise, it is fair that I should be made to hand over the equivalent

of the promised performance. In contract doctrine this proposition appears as the expectation measure of damages for breach. The expectation standard gives the victim of a breach no more or less than he would have had had there been no breach—in other words, he gets the benefit of his bargain. Two alternative measures of damage, reliance and restitution, express the different notions that if a person has relied on a promise and been hurt, that hurt must be made good; and that if a contract-breaker has obtained goods or services, he must be made to pay a fair (just?) price for them. Consider three cases:

> **2-A.** I enter your antique shop on a quiet afternoon and agree in writing to buy an expensive chest I see there, the price being about three times what you paid for it a short time ago. When I get home I repent of my decision, and within half an hour of my visit—before any other customer has come to your store—I telephone to say I no longer want the chest.
>
> **2-B.** Same as above, except in the meantime you have waxed and polished the chest and had your delivery van bring it to my door.
>
> **2-C.** Same as above, except I have the use of the chest for six months, while your shop is closed for renovations.

To require me to pay for the chest in case 2-A (or, if you resell it, to pay any profit you lost, including lost business volume) is to give you your expectation, the benefit of your bargain. In 2-B if all I must compensate is your effort I am reimbursing your reliance, and in 2-C to force me to pay a fair price for the use I have had of the chest is to focus on making me pay for, restore, an actual benefit I have received.

The assault on the classical conception of contract, the concept I call contract as promise, has centered on the connection—taken as canonical for some hundred years—between contract law and expectation damages. To focus the attack on this connection is indeed strategic. As the critics recognize and as I have just stated, to the extent that contract is grounded in promise, it seems natural to measure relief by the expectation, that is, by the promise itself. If that link can be threatened, then contract itself may be grounded elsewhere than in promise, elsewhere than in the will of the parties. In his recent comprehensive treatise, *The Rise and Fall of Freedom of Contract*, Patrick Atiyah makes the connection between the recourse to expectation damages and the emerging enforceability of executory contracts—that is, contracts enforced, though no detriment has been suffered in reliance and no benefit has been conferred. (Case 2-A is an example of an executory contract.) Before the nineteenth century, he argues, a contractual relation referred generally to one of a number of particular, community-sanctioned relations between persons who in the course of their dealings (as carriers, innkeepers, surgeons, merchants) relied on each other to their detriment or conferred benefits on each other. It was these detriments and benefits that had to be reimbursed, and an explicit promise—if there happened to be one—was important primarily to establish the reliance or to show that the benefit had been conferred in expectation of payment, not officiously or as a gift. All this, Atiyah writes, turned inside out when the promise itself came to be seen as the basis of obligation, so that neither benefit nor reliance any longer seemed necessary and the proper measure of the obligation was the promise itself, that is, the expectation. The promise principle was embraced as an expression of the principle of liberty—the will binding itself, to use Kantian language, rather than being bound by the norms of the collectivity—and the award of expectation damages followed as a natural concomitant of the promise principle.

The insistence on reliance or benefit is related to disputes about the nature of promising. As I have argued, reliance on a promise cannot alone explain its force: There is reliance because a promise is binding, and not the other way around. But if a person is bound by his promise and not by the harm the promisee may have suffered in reliance on it, then what he is bound to is just its performance. Put simply, I am bound to do what I promised you I would do—or I am bound to put you in as good a position as if I had done so. To bind me to do no more than to reimburse your reliance is to excuse me to that extent from the obligation I undertook. If your reliance is less than your expectation (in case 2-A there is no reliance), then to that extent a reliance standard excuses me from the very obligation I undertook and so weakens the force of

an obligation I chose to assume. Since by hypothesis I chose to assume the obligation in its stronger form (that is, to render the performance promised), the reliance rule indeed precludes me from incurring the very obligation I chose to undertake at the time of promising. The most compelling of the arguments for resisting this conclusion and for urging that we settle for reliance is the sense that it is sometimes harsh and ungenerous to insist on the full measure of expectancy. (This is part of Atiyah's thrust when he designates the expectation standard as an aspect of the rigid Victorian promissory morality.) The harshness comes about because in the event the promisor finds the obligation he assumed too burdensome.

This distress may be analyzed into three forms: (1) The promisor regrets having to pay for what he has bought (which may only have been the satisfaction of promising a gift or the thrill of buying a lottery ticket or stock option), though he would readily do the same thing again. I take it that this kind of regret merits no sympathy at all. Indeed if we gave in to it we would frustrate the promisor's ability to engage in his own continuing projects and so the promisor's plea is, strictly speaking, self-contradictory. (2) The promisor regrets his promise because he was mistaken about the nature of the burdens he was assuming—the purchaser in case 2-A thought he would find the money for the antique but in fact his savings are depleted, or perhaps the chest is not as old nor as valuable as he had imagined, or his house has burned down and he no longer needs it. All of these regrets are based on mistaken assumptions about the facts as they are or as they turn out to be. As we shall see . . . , the doctrines of mistake, frustration, and impossibility provide grounds for mitigating the effect of the promise principle without at all undermining it.

Finally there is the most troublesome ground of regret: (3) The promisor made no mistake about the facts or probabilities at all, but now that it has come time to perform he no longer values the promise as highly as when he made it. He regrets the promise because he regrets the value judgment that led him to make it. He concludes that the purchase of an expensive antique is an extravagance. Compassion may lead a promisee to release an obligation in such a case, but he releases as an act of generosity, not as a duty, and certainly not because the promisor's repentance destroys the force of the original obligation. The intuitive reason for holding fast is that such repentance should be the promisor's own responsibility, not one he can shift onto others. It seems too easy a way of getting out of one's obligations. Yet our intuition does not depend on suspicions of insincerity alone. Rather we feel that holding people to their obligations is a way of taking them seriously and thus of giving the concept of sincerity itself serious content. Taking this intuition to a more abstract level, I would say that respect for others as free and rational requires taking seriously their capacity to determine their own values. . . . Others must respect our capacity as free and rational persons to choose our own good, and that respect means allowing persons to take responsibility for the good they choose. . . . If we decline to take seriously the assumption of an obligation because we do not take seriously the promisor's prior conception of the good that led him to assume it, to that extent we do not take him seriously as a person. We infantilize him, as we do quite properly when we release the very young from the consequences of their choices. . . .

REMEDIES IN AND AROUND THE PROMISE

Those who have an interest in assimilating contract to the more communitarian [and utilitarian] standards of tort law have been able to obscure the link between contract and promise because in certain cases the natural thing to do *is* to give damages for the harm that has been suffered, rather than to give the money value of the promised expectation. But it does not follow from these cases that expectation is not a normal and natural measure for contract damages. First, these are situations in which the harm suffered is the measure of damages because it is hard to find the monetary value of the expectation. A leading case, *Security Stove & Mfg. Co. v. American Railway Express Co.*, illustrates the type. The plaintiff stove manufacturer had arranged to have a new kind of stove shipped by the defendant express company to a trade convention, at which the plaintiff hoped to interest prospective buyers in his improved

product. The president and his workmen went to the convention, but the defendant failed to deliver a crucial part of the exhibit in time, and they had nothing to show. Plaintiff brought suit to recover the cost of renting the booth, the freight charges, and the time and expenses lost as a result of the fruitless trip to the convention. The recovery of these items of damages, which (with the possible exception of the prepaid booth rental) seem typical examples of reliance losses, is generally agreed to have been appropriate. There was no way of knowing what results the plaintiff would have obtained had he succeeded in exhibiting his product at the convention. There was no way of knowing what his expectancy was, and so the court gave him his loss through reliance. But this illustrates only that where expectancy cannot be calculated, reliance may be a reasonable surrogate. It is reasonable to suppose that the plaintiff's expectation in *Security Stove* was at least as great as the monies he put out to exhibit his goods—after all, he was a businessman and is assumed to have been exhibiting his goods to make an eventual profit. If it could somehow be shown that the exhibit would have been a failure and the plaintiff would have suffered a net loss, the case for recovery would be undermined, and most authorities would then deny recovery.

Second are the cases in which the amount needed to undo the harm caused by reliance is itself the fairest measure of expectation.

> **3-A.** Buyer approaches manufacturer with the specifications of a small, inexpensive part—say a bolt—for a machine buyer is building. Manufacturer selects the part and sells it to buyer. The bolt is badly made, shears, and damages the machine.

The value of the thing promised, a well-made bolt, is negligible, but to give buyer his money back and no more would be a grave injustice. Here it does seem more natural to say that the manufacturer induced buyer's reasonable reliance and should compensate the resulting harm. But it is equally the case that it is a fair implication of the simple-seeming original transaction that manufacturer not only delivered and promised to transfer good title to the bolt, but promised at the same time that the bolt would do the job it was meant to do. . . .

Third, there are cases in which wrongs are committed and loss is suffered in and around the attempt to make an agreement. In these cases too reliance is the best measure of compensation. A striking example is *Hoffman v. Red Owl Stores*: A prospective Red Owl supermarket franchisee sold his previously owned business and made other expenditures on the assumption that his negotiations to obtain a Red Owl franchise would shortly be concluded. The award of reliance damages was not a case of enforcement of a promise at all, since the parties had not reached the stage where clearly determined promises had been made. Reliance damages were awarded because Red Owl had not dealt fairly with Hoffman. It had allowed him to incur expenses based on hopes that Red Owl knew or should have known were imprudent and that Red Owl was not prepared to permit him to realize. Red Owl was held liable not in order to force it to perform a promise, which it had never made, but rather to compensate Hoffman for losses he had suffered through Red Owl's inconsiderate and temporizing assurances. There is nothing at all in my conception of contract as promise that precludes persons who behave badly and cause unnecessary harm from being forced to make fair compensation. Promissory obligation is not the only basis for liability; principles of tort are sufficient to provide that people who give vague assurances that cause foreseeable harm to others should make compensation. Cases like *Hoffman* are seen to undermine the conception of contract as promise: If contract is really discrete and if it is really based in promise, then whenever there has been a promise in the picture (even only a potential promise) contractual principles must govern the whole relation. To state the argument is to reveal it as a non sequitur. It is a logical fallacy of which the classical exponents of contract as promise were themselves supremely guilty in their reluctance to grant relief for fraud or for mistakes that prevented a real agreement from coming into being. Modern critics of contractual freedom have taken the classics at their word. Justice often requires relief and adjustment in cases of accidents in and around the contracting process, and the critics have seen in this a refutation of the classics' major premise. . . . Here it is sufficient to introduce the notion that contract as promise has a distinct but neither exclusive nor necessarily dominant place among legal and moral principles.

REVIEW AND DISCUSSION QUESTIONS

1. What is the view of contract that Fried opposes? To what extent can Cohen be seen as representing that perspective?
2. Fried believes that promising is a convention but that a contract has moral force only because promises in general have force. Explain what he means. Do you agree that the obligation to keep a promise is grounded not in utility but "in respect for individual autonomy and in trust"?
3. In your view, what moral force do promises have, and why are they morally binding? What is important about the institution of promising?
4. With reference to the three versions of the antique chest story, explain the differences among expectation, reliance, and restitution as measures of damages for breach of contract. Why does Fried's view of contract lead one to the expectation measure? How do the two alternative measures of damage support a different view of the nature of contract?
5. What do you see as the pros and cons of these three measures of damages? How does Fried respond to the argument that the expectation measure is harsh and ungenerous?
6. What difficulties for Fried's view are raised by the *Security Stove* and *Hoffman* cases, and how does he respond to them?
7. Compare Fried's position on promising to that of Cohen.

Unconscionable Contracts

Williams v. Walker-Thomas Furniture Co.

Williams purchased furniture from the Walker-Thomas Furniture Company on an installment-payment plan. The terms of the contract specified that if she defaulted on her payments, the company could repossess not only the furniture in question but also any other furniture she had purchased but not yet finished paying for. After Williams defaulted, she argued that the contract was "unconscionable" and hence legally "unenforceable." The trial court and the District of Columbia Court of Appeals both ruled against her, holding that without legislative authorization they had no power to deny enforcement of contracts solely because they are unconscionable. Williams then appealed to the U.S. Circuit Court of Appeals, which held, to the contrary, that the courts can refuse to enforce such contracts, and it remanded the case to the lower court to determine whether in fact the terms of the contract were so unfair that enforcement should be withheld.

J. Skelly Wright, Circuit Judge: Appellee, Walker-Thomas Furniture Company, operates a retail furniture store in the District of Columbia. During the period from 1957 to 1962 each appellant in these cases purchased a number of household items from Walker-Thomas, for which payment was to be made in installments. The terms of each purchase were contained in a printed form contract which set forth the value of the purchased item and purported to lease the item to appellant for a stipulated monthly rent payment. The contract then provided, in substance, that title would remain in Walker-Thomas until the total of all the monthly payments made equaled the stated value of the item, at which time appellants could take title. In the event of a default in the payment of any monthly installment, Walker-Thomas could repossess the item. The contract further provided that "the amount of each periodical installment payment to be made by [purchaser] to the Company under this present lease shall be inclusive of and not in addition to the amount of each installment

350 F. 2d. 445 (D.C. Cir. 1965).

payment to be made by [purchaser] under such prior leases, bills or accounts; *and all payments now and hereafter made by [purchaser] shall be credited pro rata on all outstanding leases, bills and accounts* due the Company by [purchaser] at the time each such payment is made." (Emphasis added.) The effect of this rather obscure provision was to keep a balance due on every item purchased until the balance due on all items, whenever purchased, was liquidated. As a result, the debt incurred at the time of purchase of each item was secured by the right to repossess all the items previously purchased by the same purchaser, and each new item purchased automatically became subject to a security interest arising out of the previous dealings.

On May 12, 1962, appellant Thorne purchased an item described as a Daveno, three tables, and two lamps, having total stated value of $391.10. Shortly thereafter, he defaulted on his monthly payments and appellee sought to replevy all the items purchased since the first transaction in 1958. Similarly, on April 17, 1962, appellant Williams bought a stereo set of stated value of $514.95.[1] She too defaulted shortly thereafter, and appellee sought to replevy all the items purchased since December, 1957. The Court of General Sessions granted judgment for appellee. The District of Columbia Court of Appeals affirmed, and we granted appellants' motion for leave to appeal to this court.

Appellants' principal contention, rejected by both the trial and the appellate courts below, is that these contracts, or at least some of them, are unconscionable and, hence, not enforceable. . . . We do not agree [with the other courts] that [they] lacked the power to refuse enforcement to contracts found to be unconscionable. In other jurisdictions, it has been held as a matter of common law that unconscionable contracts are not enforceable. . . .

Unconscionability has generally been recognized to include an absence of meaningful choice on the part of one of the parties together with contract terms which are unreasonably favorable to the other party. Whether a meaningful choice is present in a particular case can only be determined by consideration of all the circumstances surrounding the transaction. In many cases the meaningfulness of the choice is negated by gross inequality of bargaining power. The manner in which the contract was entered is also relevant to this consideration. Did each party to the contract, considering his obvious education or lack of it, have a reasonable opportunity to understand the terms of the contract, or were the important terms hidden in a maze of fine print and minimized by deceptive sales practices? Ordinarily, one who signs an agreement without full knowledge of its terms might be held to assume the risk that he has entered a one-sided bargain. But when a party of little bargaining power, and hence little real choice, signs a commercially unreasonable contract with little or no knowledge of its terms, it is hardly likely that his consent, or even an objective manifestation of his consent, was ever given to all the terms. In such a case the usual rule that the terms of the agreement are not to be questioned should be abandoned and the court should consider whether the terms of the contract are so unfair that enforcement should be withheld.

In determining reasonableness or fairness, the primary concern must be with the terms of the contract considered in light of the circumstances existing when the contract was made. The test is not simple, nor can it be mechanically applied. The terms are to be considered "in the light of the general commercial background and the commercial needs of the particular trade or case."[2] . . .

Because the trial court and the appellate court did not feel that enforcement could be refused, no findings were made on the possible unconscionability of the contracts in these cases. Since the record is not sufficient for our deciding the issue as a matter of law, the cases must be remanded to the trial court for further proceedings.

So ordered.

DANAHER, Circuit Judge (dissenting): The District of Columbia Court of Appeals obviously was as unhappy about the situation here presented as any of us can possibly be. Its opinion in the Williams case . . . concludes: "We think Congress should consider corrective legislation to protect the public from such exploitive contracts as were utilized in the case at bar."

My view is thus summed up by an able court which made no finding that there had actually been sharp practice. Rather the appellant seems to have known precisely where she stood.

There are many aspects of public policy here involved. What is a luxury to some may seem an

outright necessity to others. Is public oversight to be required of the expenditures of relief funds? A washing machine, e.g., in the hands of a relief client might become a fruitful source of income. Many relief clients may well need credit, and certain business establishments will take long chances on the sale of items, expecting their pricing policies will afford a degree of protection commensurate with the risk. Perhaps a remedy when necessary will be found within the provisions of the "Loan Shark" law . . .

I mention such matters only to emphasize the desirability of a cautious approach to any such problem, particularly since the law for so long has allowed parties such great latitude in making their own contracts. I dare say there must annually be thousands upon thousands of installment credit transactions in this jurisdiction, and one can only speculate as to the effect the decision in these cases will have.

I join the District of Columbia Court of Appeals in its disposition of the issues.

NOTES

1. At the time of this purchase her account showed a balance of $164 still owing from her prior purchases. The total of all the purchases made over the years in question came to $1,800. The total payments amounted to $1,400.
2. The traditional test . . . is "such as no man in his senses and not under delusion would make on the one hand, and as no honest or fair man would accept, on the other."

REVIEW AND DISCUSSION QUESTIONS

1. According to Judge Wright, "unconscionability" requires an "absence of meaningful choice" on the part of one of the parties. What factors are relevant to determining this? In your view, was there an absence of meaningful choice in this case?
2. Unconscionability also requires "contract terms which are unreasonably favorable" to the other party. Again, how is this to be determined? In your view, were the terms of this contract "commercially unreasonable"?
3. Do you agree that, as a matter of legal principle, courts should refuse to enforce contracts they deem unconscionable but that are otherwise valid? What are the pros and cons of their doing so?
4. In his dissent Judge Danaher suggests that the furniture company was not engaging in a "sharp practice" and that Williams knew what she was doing. Explain why you agree or disagree.
5. Judge Danaher implies that poor people might be disadvantaged by the court's refusal to uphold installment credit transactions like this. Why might this be the case?
6. Judge Danaher and the lower courts believe that the enforceability of exploitative contracts is a matter that should be left to the legislature. Explain why you agree or disagree. Danaher also recommends a "cautious approach." In your view, was the majority's decision incautious?
7. How, if at all, would Cohen's and Fried's view of this case differ?

Surrogate Mother Contracts

In the Matter of Baby M

In this famous case, the Supreme Court of New Jersey ruled on the legality of surrogacy contracts: contracts, that is, in which a woman agrees, in exchange for money, to be artificially inseminated with the sperm of a stranger, to carry his child to term, and after its birth to surrender the child to the biological father. In this case, the details of which are given in the Court's opinion, William Stern sued Mary Beth

537 A.2d 1227 (1988). Citations and some notes omitted.

Whitehead for breach of contract when she refused to turn over the infant girl—called Sara by her, Melissa by the Sterns, and "Baby M" by the courts—who had been conceived with Stern's sperm. A New Jersey judge in a lower court had awarded the child to Stern and his wife, a decision later upheld by a three-judge panel. But on appeal, the New Jersey Supreme Court ruled surrogate mother contracts "illegal, perhaps criminal, and potentially degrading to women."

Chief Justice Robert Wilentz: In this matter the Court is asked to determine the validity of a contract that purports to provide a new way of bringing children into a family. For a fee of $10,000, a woman agrees to be artificially inseminated with the semen of another woman's husband; she is to conceive a child, carry it to term, and after its birth surrender it to the natural father and his wife. The intent of the contract is that the child's natural mother will thereafter be forever separated from her child. The wife is to adopt the child, and she and the natural father are to be regarded as its parents for all purposes. The contract providing for this is called a "surrogacy contract," the natural mother inappropriately called the "surrogate mother."

We invalidate the surrogacy contract because it conflicts with the law and public policy of this State. While we recognize the depth of the yearning of infertile couples to have their own children, we find the payment of money to a "surrogate" mother illegal, perhaps criminal, and potentially degrading to women. Although in this case we grant custody to the natural father, the evidence having clearly proved such custody to be in the best interests of the infant, we void both the termination of the surrogate mother's parental rights and the adoption of the child by the wife/stepparent. We thus restore the "surrogate" as the mother of the child. . . .

We find no offense to our present laws where a woman voluntarily and without payment agrees to act as a "surrogate" mother, provided that she is not subject to a binding agreement to surrender her child. Moreover, our holding today does not preclude the Legislature from altering the current statutory scheme, within constitutional limits, so as to permit surrogacy contracts. Under current law, however, the surrogacy agreement before us is illegal and invalid.

1. FACTS

In February 1985, William Stern and Mary Beth Whitehead entered into a surrogacy contract. It recited that Stern's wife, Elizabeth, was infertile, that they wanted a child, and that Mrs. Whitehead was willing to provide that child as the mother with Mr. Stern as the father.

The contract provided that through artificial insemination using Mr. Stern's sperm, Mrs. Whitehead would become pregnant, carry the child to term, bear it, deliver it to the Sterns, and thereafter do whatever was necessary to terminate her maternal rights so that Mrs. Stern could thereafter adopt the child. Mrs. Whitehead's husband, Richard,[1] was also a party to the contract; Mrs. Stern was not. . . .

Mr. Stern, on his part, agreed to attempt the artificial insemination and to pay Mrs. Whitehead $10,000 after the child's birth, on its delivery to him. In a separate contract, Mr. Stern agreed to pay $7,500 to the Infertility Center of New York ("ICNY"). The Center's advertising campaigns solicit surrogate mothers and encourage infertile couples to consider surrogacy. ICNY arranged for the surrogacy contract by bringing the parties together, explaining the process to them, furnishing the contractual form, and providing legal counsel.

The history of the parties' involvement in this arrangement suggests their good faith. William and Elizabeth Stern were married in July 1974, having met at the University of Michigan, where both were Ph.D. candidates. Due to financial considerations and Mrs. Stern's pursuit of a medical degree and residency, they decided to defer starting a family until 1981. Before then, however, Mrs. Stern learned that she might have multiple sclerosis and that the disease in some cases renders pregnancy a serious health risk. Her anxiety appears to have exceeded the actual risk, which current medical authorities assess as minimal. Nonetheless that anxiety was evidently quite real, Mrs. Stern fearing that pregnancy might precipitate blindness, paraplegia, or other forms of debilitation. Based on the perceived risk, the Sterns decided to forego having their own children. The decision had a special significance for Mr. Stern. Most of his family had been destroyed in the Holocaust. As the family's only survivor, he very much wanted to continue his bloodline. . . .

On February 6, 1985, Mr. Stern and Mr. and Mrs. Whitehead executed the surrogate parenting agreement. After several artificial inseminations over a period of months, Mrs. Whitehead became pregnant. The pregnancy was uneventful and on March 27, 1986, Baby M was born.

Not wishing anyone at the hospital to be aware of the surrogacy arrangement, Mr. and Mrs. Whitehead appeared to all as the proud parents of a healthy female child. Her birth certificate indicated her name to be Sara Elizabeth Whitehead and her father to be Richard Whitehead. In accordance with Mrs. Whitehead's request, the Sterns visited the hospital unobtrusively to see the newborn child.

Mrs. Whitehead realized, almost from the moment of birth, that she could not part with this child. She had felt a bond with it even during pregnancy. Some indication of the attachment was conveyed to the Sterns at the hospital when they told Mrs. Whitehead what they were going to name the baby. She apparently broke into tears and indicated that she did not know if she could give up the child. She talked about how the baby looked like her other daughter, and made it clear that she was experiencing great difficulty with the decision.

Nonetheless, Mrs. Whitehead was, for the moment, true to her word. Despite powerful inclinations to the contrary, she turned her child over to the Sterns on March 30 at the Whitehead's home. . . .

Later in the evening of March 30, Mrs. Whitehead became deeply disturbed, disconsolate, stricken with unbearable sadness. She had to have her child. She could not eat, sleep, or concentrate on anything other than her need for her baby. The next day she went to the Sterns' home and told them how much she was suffering.

The depth of Mrs. Whitehead's despair surprised and frightened the Sterns. She told them that she could not live without her baby, that she must have her, even if only for one week, that thereafter she would surrender her child. The Sterns, concerned that Mrs. Whitehead might indeed commit suicide, not wanting under any circumstances to risk that, and in any event believing that Mrs. Whitehead would keep her word, turned the child over to her. It was not until four months later, after a series of attempts to regain possession of the child, that Melissa was returned to the Sterns, having been forcibly removed from the home where she was then living with Mr. and Mrs. Whitehead, the home in Florida owned by Mary Beth Whitehead's parents.

The struggle over Baby M began when it became apparent that Mrs. Whitehead could not return the child to Mr. Stern. Due to Mrs. Whitehead's refusal to relinquish the baby, Mr. Stern filed a complaint seeking enforcement of the surrogacy contract. He alleged, accurately, that Mrs. Whitehead had not only refused to comply with the surrogacy contract but had threatened to flee from New Jersey with the child in order to avoid even the possibility of his obtaining custody. The court papers asserted that if Mrs. Whitehead were to be given notice of the application for an order requiring her to relinquish custody, she would, prior to the hearing, leave the state with the baby. And that is precisely what she did. After the order was entered, *ex parte*, the process server, aided by the police, in the presence of the Sterns, entered Mrs. Whitehead's home to execute the order. Mr. Whitehead fled with the child, who had been handed to him through a window while those who came to enforce the order were thrown off balance by a dispute over the child's current name.

The Whiteheads immediately fled to Florida with Baby M. They stayed initially with Mrs. Whitehead's parents, where one of Mrs. Whitehead's children had been living. For the next three months, the Whiteheads and Melissa lived at roughly twenty different hotels, motels, and homes in order to avoid apprehension. From time to time Mrs. Whitehead would call Mr. Stern to discuss the matter; the conversations, recorded by Mr. Stern on advice of counsel, show an escalating dispute about rights, morality, and power, accompanied by threats of Mrs. Whitehead to kill herself, to kill the child, and falsely to accuse Mr. Stern of sexually molesting Mrs. Whitehead's other daughter.

Eventually the Sterns discovered where the Whiteheads were staying, commenced supplementary proceedings in Florida, and obtained an order requiring the Whiteheads to turn over the child. Police in Florida enforced the order, forcibly removing the child from her grandparents' home. She was soon thereafter brought to New Jersey and turned over to the Sterns. . . .

The trial took thirty-two days over a period of more than two months. It included numerous interlocutory appeals and attempted interlocutory

appeals. There were twenty-three witnesses to the facts recited above and fifteen expert witnesses, eleven testifying on the issue of custody and four on the subject of Mrs. Stern's multiple sclerosis; the bulk of the testimony was devoted to determining the parenting arrangement most compatible with the child's best interests. Soon after the conclusion of the trial, the trial court announced its opinion from the bench. It held that the surrogacy contract was valid; ordered that Mrs. Whitehead's parental rights be terminated and that sole custody of the child be granted to Mr. Stern; and, after hearing brief testimony from Mrs. Stern, immediately entered an order allowing the adoption of Melissa by Mrs. Stern, all in accordance with the surrogacy contract. Pending the outcome of the appeal, we granted a continuation of visitation to Mrs. Whitehead, although slightly more limited than the visitation allowed during the trial. . . .

Mrs. Whitehead appealed. . . .

Mrs. Whitehead contends that the surrogacy contract, for a variety of reasons, is invalid. She contends that it conflicts with public policy since it guarantees that the child will not have the nurturing of both natural parents—presumably New Jersey's goal for families. She further argues that it deprives the mother of her constitutional right to the companionship of her child, and that it conflicts with statutes concerning termination of parental rights and adoption. With the contract thus void, Mrs. Whitehead claims primary custody (with visitation rights in Mr. Stern) both on a best interests basis (stressing the "tender years" doctrine) as well as the policy basis of discouraging surrogacy contracts. She maintains that even if custody would ordinarily go to Mr. Stern, here it should be awarded to Mrs. Whitehead to deter future surrogacy arrangements. . . .

2. INVALIDITY AND UNENFORCEABILITY OF SURROGACY CONTRACT

We have concluded that this surrogacy contract is invalid. Our conclusion has two bases: direct conflict with existing statutes and conflict with the public policies of this State, as expressed in its statutory and decisional law.

One of the surrogacy contract's basic purposes, to achieve the adoption of a child through private placement, though permitted in New Jersey "is very much disfavored." *Sees v. Baber.* Its use of money for this purpose—and we have no doubt whatsoever that the money is being paid to obtain an adoption and not, as the Sterns argue, for the personal services of Mary Beth Whitehead—is illegal and perhaps criminal. In addition to the inducement of money, there is the coercion of contract: the natural mother's irrevocable agreement, prior to birth, even prior to conception, to surrender the child to the adoptive couple. Such an agreement is totally unenforceable in private placement adoption. Even where the adoption is through an approved agency, the formal agreement to surrender occurs only *after* birth (as we read N.J.S.A. 9:2–16 and 17, and similar statutes), and then, by regulation, only after the birth mother has been counseled. . . .

A. Conflict with Statutory Provisions

The surrogacy contract conflicts with: . . . laws prohibiting the use of money in connection with adoptions . . . [and] laws that make surrender of custody and consent to adoption revocable in private placement adoptions.

Our law prohibits paying or accepting money in connection with any placement of a child for adoption. Violation is a high misdemeanor. Excepted are fees of an approved agency (which must be a non-profit entity) and certain expenses in connection with childbirth.

Considerable care was taken in this case to structure the surrogacy arrangement so as not to violate this prohibition. The arrangement was structured as follows: the adopting parent, Mrs. Stern, was not a party to the surrogacy contract; the money paid to Mrs. Whitehead was stated to be for her services—not for the adoption; the sole purpose of the contract was stated as being that "of giving a child to William Stern, its natural and biological father"; the money was purported to be "compensation for services and expenses and in no way . . . a fee for termination of parental rights or a payment in exchange for consent to surrender a child for adoption"; the fee to the Infertility Center ($7,500) was stated to be for legal representation, advice, administrative work, and other "services." Nevertheless, it seems clear that the

money was paid and accepted in connection with an adoption. . . .

Mr. Stern knew he was paying for the adoption of a child; Mrs. Whitehead knew she was accepting money so that a child might be adopted; the Infertility Center knew that it was being paid for assisting in the adoption of a child. The actions of all three worked to frustrate the goals of the statute. It strains credulity to claim that these arrangements, touted by those in the surrogacy business as an attractive alternative to the usual route leading to an adoption, really amount to something other than a private placement adoption for money.

The prohibition of our statute is strong. Violation constitutes a high misdemeanor, a third-degree crime, carrying a penalty of three to five years imprisonment. The evils inherent in baby bartering are loathsome for a myriad of reasons. The child is sold without regard for whether the purchasers will be suitable parents. N. Baker, *Baby Selling: The Scandal of Black Market Adoption* (1978). The natural mother does not receive the benefit of counseling and guidance to assist her in making a decision that may affect her for a lifetime. In fact, the monetary incentive to sell her child may, depending on her financial circumstances, make her decision less voluntary. Furthermore, the adoptive parents may not be fully informed of the natural parents' medical history.

Baby-selling potentially results in the exploitation of all parties involved. Conversely, adoption statutes seek to further humanitarian goals, foremost among them the best interests of the child. H. Witmer, E. Herzog, E. Weinstein, & M. Sullivan, *Independent Adoptions: A Follow-Up Study* 32 (1967). The negative consequences of baby buying are potentially present in the surrogacy context, especially the potential for placing and adopting a child without regard to the interest of the child or the natural mother. . . .

The provision in the surrogacy contract stating that Mary Beth Whitehead agrees to "surrender custody . . . and terminate all parental rights" contains no clause giving her a right to rescind. It is intended to be an irrevocable consent to surrender the child for adoption—in other words, an irrevocable commitment by Mrs. Whitehead to turn Baby M over to the Sterns and thereafter to allow termination of her parental rights. The trial court required a "best interests" showing as a condition to granting specific performance of the surrogacy contract. Having decided the "best interests" issue in favor of the Sterns, that court's order included, among other things, specific performance of this agreement to surrender custody and terminate all parental rights.

Mrs. Whitehead, shortly after the child's birth, had attempted to revoke her consent and surrender by refusing, after the Sterns had allowed her to have the child "just for one week," to return Baby M to them. The trial court's award of specific performance therefore reflects its view that the consent to surrender the child was irrevocable. We accept the trial court's construction of the contract; indeed it appears quite clear that this was the parties' intent. Such a provision, however, making irrevocable the natural mother's consent to surrender custody of her child in a private placement adoption, clearly conflicts with New Jersey law.

Our analysis commences with the statute providing for surrender of custody to an approved agency and termination of parental rights on the suit of that agency. The two basic provisions of the statute are N.J.S.A. 9:2–14 and 9:2–16. The former provides explicitly that:

> Except as otherwise provided by law or by order or judgment of a court of competent jurisdiction or by testamentary disposition, no surrender of the custody of a child shall be valid in this state unless made to an approved agency pursuant to the provisions of this act. . . .

. . . Requirements for a voluntary surrender to an approved agency was set forth in N.J.S.A. 9:2–16. This section allows an approved agency to take a voluntary surrender of custody from the parent of a child but provides stringent requirements as a condition to its validity.

. . . These strict prerequisites to irrevocability constitute a recognition of the most serious consequences that flow from such consents: termination of parental rights, the permanent separation of parent from child, and the ultimate adoption of the child. Because of those consequences, the Legislature severely limited the circumstances under which such consent would be irrevocable. The legislative goal is furthered by regulations requiring approved agencies, prior to accepting irrevocable consents, to provide advice and

counseling to women, making it more likely that they fully understand and appreciate the consequences of their acts.

Contractual surrender of parental rights is not provided for in our statutes as now written. . . . There is no doubt that a contractual provision purporting to constitute an irrevocable agreement to surrender custody of a child for adoption is invalid. . . .

B. Public Policy Considerations

The surrogacy contract's invalidity, resulting from its direct conflict with the above statutory provisions, is further underlined when its goals and means are measured against New Jersey's public policy. The contract's basic premise, that the natural parents can decide in advance of birth which one is to have custody of the child, bears no relationship to the settled law that the child's best interests shall determine custody. . . .

The surrogacy contract guarantees permanent separation of the child from one of its natural parents. Our policy, however, has long been that to the extent possible, children should remain with and be brought up by both of their natural parents. . . . This is not simply some theoretical ideal that in practice has no meaning. The impact of failure to follow that policy is nowhere better shown than in the results of this surrogacy contract. A child, instead of starting off its life with as much peace and security as possible, finds itself immediately in a tug-of-war between contending mother and father.

The surrogacy contract violates the policy of this State that the rights of natural parents are equal concerning their child, the father's right no greater than the mother's. "The parent and child relationship extends equally to every child and to every parent, regardless of the marital status of the parents." . . . The whole purpose and effect of the surrogacy contract was to give the father the exclusive right to the child by destroying the rights of the mother.

The policies expressed in our comprehensive laws governing consent to the surrender of a child, stand in stark contrast to the surrogacy contract and what it implies. Here there is no counseling, independent or otherwise, of the natural mother, no evaluation, no warning.

The only legal advice Mary Beth Whitehead received regarding the surrogacy contract was provided in connection with the contract that she previously entered into with another couple. Mrs. Whitehead's lawyer was referred to her by the Infertility Center, with which he had an agreement to act as counsel for surrogate candidates. His services consisted of spending one hour going through the contract with the Whiteheads, section by section, and answering their questions. Mrs. Whitehead received no further legal advice prior to signing the contract with the Sterns.

Mrs. Whitehead was examined and psychologically evaluated, but if it was for her benefit, the record does not disclose that fact. The Sterns regarded the evaluation as important, particularly in connection with the question of whether she would change her mind. Yet they never asked to see it, and were content with the assumption that the Infertility Center had made an evaluation and had concluded that there was no danger that the surrogate mother would change her mind. From Mrs. Whitehead's point of view, all that she learned from the evaluation was that "she had passed." It is apparent that the profit motive got the better of the Infertility Center. Although the evaluation was made, it was not put to any use, and understandably so, for the psychologist warned that Mrs. Whitehead demonstrated certain traits that might make surrender of the child difficult and that there should be further inquiry into this issue in connection with her surrogacy. To inquire further, however, might have jeopardized the Infertility Center's fee. The record indicates that neither Mrs. Whitehead nor the Sterns were ever told of this fact, a fact that might have ended their surrogacy arrangement.

Under the contract, the natural mother is irrevocably committed before she knows the strength of her bond with her child. She never makes a totally voluntary, informed decision, for quite clearly any decision prior to the baby's birth is, in the most important sense, uninformed, and any decision after that, compelled by a pre-existing contractual commitment, the threat of a lawsuit, and the inducement of a $10,000 payment, is less than totally voluntary. Her interests are of little concern to those who controlled this transaction.

Although the interest of the natural father and adoptive mother is certainly the predominant interest, realistically the *only* interest served, even they are left with less than what public policy requires. They know little about the natural mother, her genetic makeup, and her psychological and medical history. Moreover, not even a superficial attempt is made to determine their awareness of their responsibilities as parents.

Worst of all, however, is the contract's total disregard of the best interests of the child. There is not the slightest suggestion that any inquiry will be made at any time to determine the fitness of the Sterns as custodial parents, of Mrs. Stern as an adoptive parent, their superiority to Mrs. Whitehead, or the effect on the child of not living with her natural mother.

This is the sale of a child, or, at the very least, the sale of a mother's right to her child, the only mitigating factor being that one of the purchasers is the father. Almost every evil that prompted the prohibition of the payment of money in connection with adoptions exists here.

The differences between an adoption and a surrogacy contract should be noted, since it is asserted that the use of money in connection with surrogacy does not pose the risks found where money buys an adoption.

First, and perhaps most important, all parties concede that it is unlikely that surrogacy will survive without money. Despite the alleged selfless motivation of surrogate mothers, if there is no payment, there will be no surrogates, or very few. That conclusion contrasts with adoption; for obvious reasons, there remains a steady supply, albeit insufficient, despite the prohibitions against payment. The adoption itself, relieving the natural mother of the financial burden of supporting an infant, is the equivalent of payment.

Second, the use of money in adoptions does not *produce* the problem—conception occurs, and usually the birth itself, before illicit funds are offered. With surrogacy, the "problem," if one views it as such, consisting of the purchase of a woman's procreative capacity, at the risk of her life, is caused by and originates with the offer of money.

Third, with the law prohibiting the use of money in connection with adoptions, the built-in financial pressure of the unwanted pregnancy and the consequent support obligation do not lead the mother to the highest paying, ill-suited, adoptive parents. She is just as well off surrendering the child to an approved agency. In surrogacy, the highest bidders will presumably become the adoptive parents regardless of suitability, so long as payment of money is permitted.

Fourth, the mother's consent to surrender her child in adoptions is revocable, even after surrender of the child, unless it be to an approved agency, where by regulation there are protections against an ill-advised surrender. In surrogacy, consent occurs so early that no amount of advice would satisfy the potential mother's need, yet the consent is irrevocable.

The main difference, that the plight of the unwanted pregnancy is unintended while the situation of the surrogate mother is voluntary and intended, is really not significant. Initially, it produces stronger reactions of sympathy for the mother whose pregnancy was unwanted than for the surrogate mother, who "went into this with her eyes wide open." On reflection, however, it appears that the essential evil is the same, taking advantage of a woman's circumstances (the unwanted pregnancy or the need for money) in order to take away her child, the difference being one of degree. . . .

Intimated, but disputed, is the assertion that surrogacy will be used for the benefit of the rich at the expense of the poor. In response it is noted that the Sterns are not rich and the Whiteheads not poor. Nevertheless, it is clear to us that it is unlikely that surrogate mothers will be as proportionately numerous among those women in the top twenty percent income bracket as among those in the bottom twenty percent. Put differently, we doubt that infertile couples in the low-income bracket will find upper income surrogates.

In any event, even in this case one should not pretend that disparate wealth does not play a part simply because the contrast is not the dramatic "rich versus poor." At the time of trial, the Whiteheads' net assets were probably negative—Mrs. Whitehead's own sister was foreclosing on a second mortgage. Their income derived from Mr. Whitehead's labors. Mrs. Whitehead is a homemaker, having previously held part-time jobs. The Sterns are both professionals, she a medical doctor, he a biochemist. Their combined income when both were working was about $89,500 a year and

their assets sufficient to pay for the surrogacy contract arrangements.

The point is made that Mrs. Whitehead *agreed* to the surrogacy arrangement, supposedly fully understanding the consequences. Putting aside the issue of how compelling her need for money may have been, and how significant her understanding of the consequences, we suggest that her consent is irrelevant. There are, in a civilized society, some things that money cannot buy. In America, we decided long ago that merely because conduct purchased by money was "voluntary" did not mean that it was good or beyond regulation and prohibition. Employers can no longer buy labor at the lowest price they can bargain for, even though that labor is "voluntary," or buy women's labor for less money than paid to men for the same job, or purchase the agreement of children to perform oppressive labor, or purchase the agreement of workers to subject themselves to unsafe or unhealthy working conditions (Occupational Health and Safety Act of 1970). There are, in short, values that society deems more important than granting to wealth whatever it can buy, be it labor, love, or life. Whether this principle recommends prohibition of surrogacy, which presumably sometimes results in great satisfaction to all of the parties, is not for us to say. We note here only that, under existing law, the fact that Mrs. Whitehead "agreed" to the arrangement is not dispositive.

The long-term effects of surrogacy contracts are not known, but feared—the impact on the child who learns her life was bought, that she is the offspring of someone who gave birth to her only to obtain money; the impact on the natural mother as the full weight of her isolation is felt along with the full reality of the sale of her body and her child; the impact on the natural father and adoptive mother once they realize the consequences of their conduct. Literature in related areas suggests these are substantial considerations, although, given the newness of surrogacy, there is little information.

The surrogacy contract creates, [and] it is based upon, principles that are directly contrary to the objectives of our laws. It guarantees the separation of a child from its mother; it looks to adoption regardless of suitability; it totally ignores the child; it takes the child from the mother regardless of her wishes and her maternal fitness; and it does all of this, it accomplishes all of its goals, through the use of money.

Beyond that is the potential degradation of some women that may result from this arrangement. In many cases, of course, surrogacy may bring satisfaction, not only to the infertile couple, but to the surrogate mother herself. The fact, however, that many women may not perceive surrogacy negatively but rather see it as an opportunity does not diminish its potential for devastation to other women.

In sum, the harmful consequences of this surrogacy arrangement appear to us all too palpable. In New Jersey the surrogate mother's agreement to sell her child is void. Its irrevocability infects the entire contract, as does the money that purports to buy it. . . .

CONCLUSION

This case affords some insight into a new reproductive arrangement: the artificial insemination of a surrogate mother. The unfortunate events that have unfolded illustrate that its unregulated use can bring suffering to all involved. Potential victims include the surrogate mother and her family, the natural father and his wife, and most importantly, the child. Although surrogacy has apparently provided positive results for some infertile couples, it can also, as this case demonstrates, cause suffering to participants, here essentially innocent and well intended.

We have found that our present laws do not permit the surrogacy contract used in this case. Nowhere, however, do we find any legal prohibition against surrogacy when the surrogate mother volunteers, without any payment, to act as a surrogate and is given the right to change her mind and to assert her parental rights. Moreover, the Legislature remains free to deal with this most sensitive issue as it sees fit, subject only to constitutional constraints.

If the Legislature decides to address surrogacy, consideration of this case will highlight many of its potential harms. We do not underestimate the difficulties of legislating on this subject. In addition to the inevitable confrontation with the ethical and moral issues involved, there is the question of the wisdom and effectiveness of regulating a matter so private, yet of such public interest. Legislative consideration of surrogacy may also provide the

opportunity to begin to focus on the overall implications of the new reproductive biotechnology—*in vitro* fertilization, preservation of sperm and eggs, embryo implantation and the like. The problem is how to enjoy the benefits of the technology—especially for infertile couples—while minimizing the risk of abuse. The problem can be addressed only when society decides what its values and objectives are in this troubling, yet promising, area.

The judgment is affirmed in part, reversed in part, and remanded for further proceedings consistent with this opinion.

NOTE

1. Subsequent to the trial court proceedings, Mr. and Mrs. Whitehead were divorced, and soon thereafter Mrs. Whitehead remarried. Nevertheless, in the course of this opinion we will make reference almost exclusively to the facts as they existed at the time of trial, the facts on which the decision we now review was reached. We note moreover that Mr. Whitehead remains a party to this dispute. For these reasons, we continue to refer to appellants as Mr. and Mrs. Whitehead.

REVIEW AND DISCUSSION QUESTIONS

1. Justice Wilentz contends that surrogacy contracts conflict with New Jersey statute. Why? Do you think Wilentz would have ruled that there was a conflict if he had approved of such contracts?
2. Is the Baby M case a case of "baby selling," as Justice Wilentz implies, or are surrogate agencies correct to claim that they are simply selling a woman's services? Under either description, is there anything morally questionable about surrogate mother agreements?
3. What public policy considerations does Justice Wilentz raise against surrogacy contracts? What public policy considerations might be raised in their favor? Did Justice Wilentz make the correct decision?
4. Assess this case from Fried's theory of contract as promise. Does Wilentz's releasing Mary Beth Whitehead from her contract with Stern "infantilize" her (in Fried's words) and fail to take her "seriously as a person"? Or are such contracts, as Wilentz suggests, degrading to women? Do you think Justice Wilentz's view of contract law is closer to Cohen's or to Fried's?
5. Could surrogacy contracts be modified or regulated in some way that gets around the objections that Justice Wilentz and other critics have of them?

Employment Contracts

Lochner v. New York; Muller v. Oregon; Coppage v. Kansas; and *West Coast Hotel Co. v. Parrish*

Although common law establishes the general right of individuals to make employment contracts on any mutually agreeable terms, the courts have long permitted the government to restrict certain sorts of contracts when necessary to promote the safety, health, morals, or general welfare of the public. In the famous *Lochner* case of 1905, however, the U.S. Supreme Court struck down a New York labor law that prevented bakers from working more than sixty hours a week on the ground that the statute unreasonably interfered with the right of contract between employer and employees. Three years later, though, it upheld an Oregon statute that prevented women from working more than ten hours a day in factories and laundries. In 1915, in the *Coppage* case, the Court reaffirmed its *Lochner* position by invalidating a

198 U.S. 45 (1905), 208 U.S. 412 (1908), 236 U.S. 1 (1915), and 300 U.S. 379 (1937).

Kansas law that forbade "yellow dog" contracts, that is, contracts in which employers require, as a condition of employment, that employees not belong to a labor union. In the early 1930s, the Supreme Court's resistance to government interference with employment contracts brought it into collision with President Franklin D. Roosevelt, who threatened to add new members to the Court who would uphold his progressive labor legislation. Thereafter, the Court backed away from its position in *Lochner* and *Coppage*. In *West Coast Hotel Co. v. Parrish*, the Supreme Court reversed its own position in an earlier case, explicitly upholding a minimum-wage law and implicitly removing any constitutional barriers to the regulation of labor contracts.

Lochner v. New York

Mr. Justice Peckham delivered the opinion of the Court: The indictment, it will be seen, charges that the plaintiff in error violated the 110th section of article 8, chapter 415, of the Laws of 1897, known as the labor law of the state of New York, in that he wrongfully and unlawfully required and permitted an employee working for him to work more than sixty hours in one week. . . .

It is not an act merely fixing the number of hours which shall constitute a legal day's work, but an absolute prohibition upon the employer permitting, under any circumstances, more than ten hours' work to be done in his establishment. The employee may desire to earn the extra money which would arise from his working more than the prescribed time, but this statute forbids the employer from permitting the employee to earn it.

The statute necessarily interferes with the right of contract between the employer and employees, concerning the number of hours in which the latter may labor in the bakery of the employer. The general right to make a contract in relation to his business is part of the liberty of the individual protected by the 14th Amendment of the Federal Constitution. . . . Under that provision no state can deprive any person of life, liberty, or property without due process of law. The right to purchase or to sell labor is part of the liberty protected by this amendment, unless there are circumstances which exclude the right. There are, however, certain powers, existing in the sovereignty of each state in the Union, somewhat vaguely termed police powers, the exact description and limitation of which have not been attempted by the courts. Those powers, broadly stated, and without, at present, any attempt at a more specific limitation, relate to the safety, health, morals, and general welfare of the public. Both property and liberty are held on such reasonable conditions as may be imposed by the governing power of the state in the exercise of those powers, and with such conditions the 14th Amendment was not designed to interfere. . . .

Therefore, when the state, by its legislature, in the assumed exercise of its police powers, has passed an act which seriously limits the right to labor or the right of contract in regard to their means of livelihood between persons who are *sui juris* (both employer and employee), it becomes of great importance to determine which shall prevail—the right of the individual to labor for such time as he may choose, or the right of the state to prevent the individual from laboring, or from entering into any contract to labor, beyond a certain time prescribed by the state.

This court has recognized the existence and upheld the exercise of the police powers of the states in many cases which might fairly be considered as border ones, and it has, in the course of its determination of questions regarding the asserted invalidity of such statutes, on the ground of their violation of the rights secured by the Federal Constitution, been guided by rules of a very liberal nature, the application of which has resulted, in numerous instances, in upholding the validity of state statutes thus assailed. Among the later cases where the state law has been upheld by this court is that of *Holden v. Hardy*, 169 U.S. 366. A provision in the act of the legislature of Utah was there under consideration, the act limiting the employment of workmen in all underground mines or workings, to eight hours per day, "except in cases of emergency, where life or property is in imminent danger." It also limited the hours of labor in smelting and other institutions for the reduction or refining of

ores or metals to eight hours per day, except in like cases of emergency. The act was held to be a valid exercise of the police powers of the state. . . .

It must, of course, be conceded that there is a limit to the valid exercise of the police power by the state. . . .

In every case that comes before this court, therefore, where legislation of this character is concerned, and where the protection of the Federal Constitution is sought, the question necessarily arises: Is this a fair, reasonable, and appropriate exercise of the police power of the state, or is it an unreasonable, unnecessary, and arbitrary interference with the right of the individual to his personal liberty, or to enter into those contracts in relation to labor which may seem to him appropriate or necessary for the support of himself and his family? Of course the liberty of contract relating to labor includes both parties to it. The one has as much right to purchase as the other to sell labor.

This is not a question of substituting the judgment of the court for that of the legislature. If the act be within the power of the state it is valid, although the judgment of the court might be totally opposed to the enactment of such a law. But the question would still remain: Is it within the police power of the state? and that question must be answered by the court.

The question whether this act is valid as a labor law, pure and simple, may be dismissed in a few words. There is no reasonable ground for interfering with the liberty of persons or the right of free contract by determining the hours of labor, in the occupation of a baker. . . . Viewed in the light of a purely labor law with no reference whatever to the question of health, we think that a law like the one before us involves neither the safety, the morals, nor the welfare of the public, and that the interest of the public is not in the slightest degree affected by such an act. The law must be upheld, if at all, as a law pertaining to the health of the individual engaged in the occupation of a baker. . . .

We think the limit of the police power has been reached and passed in this case. There is, in our judgment, no reasonable foundation for holding this to be necessary or appropriate as a health law to safeguard the public health, or the health of the individuals who are following the trade of a baker. . . .

It is also urged, pursuing the same line of argument, that it is to the interest of the state that its population should be strong and robust, and therefore any legislation which may be said to tend to make people healthy must be valid as health laws, enacted under the police power. If this be a valid argument and a justification for this kind of legislation, it follows that the protection of the Federal Constitution from undue interference with liberty of person and freedom of contract is visionary, wherever the law is sought to be justified as a valid exercise of the police power. Scarcely any law but might find shelter under such assumptions, and conduct, properly so called, as well as contract, would come under the restrictive sway of the legislature. Not only the hours of employees, but the hours of employers, could be regulated, and doctors, lawyers, scientists, all professional men, as well as athletes and artisans, could be forbidden to fatigue their brains and bodies by prolonged hours of exercise, lest the fighting strength of the state be impaired. We mention these extreme cases because the contention is extreme. We do not believe in the soundness of the views which uphold this law. . . .

Statutes of the nature of that under review, limiting the hours in which grown and intelligent men may labor to earn their living, are mere meddlesome interferences with the rights of the individual, and they are not saved from condemnation by the claim that they are passed in the exercise of the police power and upon the subject of the health of the individual whose rights are interfered with, unless there be some fair ground, reasonable in and of itself, to say that there is material danger to the public health, or to the health of the employees, if the hours of labor are not curtailed. . . .

It is impossible for us to shut our eyes to the fact that many of the laws of this character, while passed under what is claimed to be the police power for the purpose of protecting the public health or welfare, are, in reality, passed from other motives. We are justified in saying so when, from the character of the law and the subject upon which it legislates, it is apparent that the public health or welfare bears but the most remote relation to the law. The purpose of a statute must be determined from the natural and legal effect of the language employed; and whether it is or is not repugnant to the Constitution of the United States must be determined from the natural

effect of such statutes when put into operation, and not from their proclaimed purpose. . . .

It is manifest to us that the limitation of the hours of labor as provided for in this section of the statute under which the indictment was found, and the plaintiff in error convicted, has no such direct relation to, and no such substantial effect upon, the health of the employee, as to justify us in regarding the section as really a health law. It seems to us that the real object and purpose were simply to regulate the hours of labor between the master and his employees (all being men, *sui juris*), in a private business, not dangerous in any degree to morals, or in any real and substantial degree to the health of the employees. Under such circumstances the freedom of master and employee to contract with each other in relation to their employment, and in defining the same, cannot be prohibited or interfered with, without violating the Federal Constitution.

The judgment of the Court of Appeals of New York, as well as that of the Supreme Court and of the County Court of Oneida County, must be reversed and the case remanded to the County Court for further proceedings not inconsistent with this opinion.

Reversed.

Mr. Justice Holmes, dissenting: I regret sincerely that I am unable to agree with the judgment in this case, and that I think it my duty to express my dissent.

This case is decided upon an economic theory which a large part of the country does not entertain. If it were a question whether I agreed with that theory, I should desire to study it further and long before making up my mind. But I do not conceive that to be my duty, because I strongly believe that my agreement or disagreement has nothing to do with the right of a majority to embody their opinions in law. It is settled by various decisions of this court that state constitutions and state laws may regulate life in many ways which we as legislators might think as injudicious, or if you like as tyrannical, as this, and which, equally with this, interfere with the liberty to contract. Sunday laws and usury laws are ancient examples. A more modern one is the prohibition of lotteries. The liberty of the citizen to do as he likes so long as he does not interfere with the liberty of others to do the same which has been a shibboleth for some well-known writers, is interfered with by school laws, by the Postoffice, by every state or municipal institution which takes his money for purposes though desirable, whether he likes it or not. The 14th Amendment does not enact Mr. Herbert Spencer's *Social Statics*. The other day we sustained the Massachusetts vaccination law. . . . United States and state statutes and decisions cutting down the liberty to contract by way of combination are familiar to this court. . . . Two years ago we upheld the prohibition of sales of stock on margins, or for future delivery, in the Constitution of California. . . . The decision sustaining an eight-hour law for miners is still recent. . . . Some of these laws embody convictions or prejudices which judges are likely to share. Some may not. But a Constitution is not intended to embody a particular economic theory, whether of paternalism and the organic relation of the citizen to the state or of *laissez faire*. It is made for people of fundamentally differing views, and the accident of our finding certain opinions natural and familiar, or novel, and even shocking, ought not to conclude our judgment upon the question whether statutes embodying them conflict with the Constitution of the United States.

General propositions do not decide concrete cases. The decision will depend on a judgment or intuition more subtle than any articulate major premise. But I think that the proposition just stated, if it is accepted, will carry us far toward the end. Every opinion tends to become a law. I think that the word "liberty," in the 14th Amendment, is perverted when it is held to prevent the natural outcome of a dominant opinion, unless it can be said that a rational and fair man necessarily would admit that the statute proposed would infringe fundamental principles as they have been understood by the traditions of our people and our law. It does not need research to show that no such sweeping condemnation can be passed upon the statute before us. A reasonable man might think it a proper measure on the score of health. Men whom I certainly could not pronounce unreasonable would uphold it as a first installment of a general regulation of the hours of work. Whether in the latter aspect it would be open to the charge of inequality I think it unnecessary to discuss.

Mr. Justice Harlan (with whom Mr. Justice White and Mr. Justice Day concurred) also dissented.

Muller v. Oregon

In this case, the Court upheld an Oregon statute that prevented women from working in factories or laundries for more than ten hours a day. In doing so, it relied on a brief by Louis Brandeis that distinguished the present case from the *Lochner* precedent on the grounds that women are physically different from men. Thus, the Court interpreted the Oregon statue as health regulation properly within the police power of the state.

In patent cases counsel are apt to open the argument with a discussion of the state of the art. It may not be amiss, in the present case, before examining the constitutional question, to notice the course of legislation as well as expressions of opinion from other than judicial sources. In the brief filed by Mr. Louis D. Brandeis, for the defendant in error, is a very copious collection of all these matters, an epitome of which is found in the margin. . . .

The legislation and opinions referred to in the margin may not be, technically speaking, authorities, and in them is little or no discussion of the constitutional question presented to us for determination, yet they are significant of a widespread belief that woman's physical structure, and the functions she performs in consequence thereof, justify special legislation restricting or qualifying the conditions under which she should be permitted to toil. Constitutional questions, it is true, are not settled by even a consensus of present public opinion, for it is the peculiar value of a written constitution that it places in unchanging form limitations upon legislative action, and thus gives a permanence and stability to popular government which otherwise would be lacking. At the same time, when a question of fact is debated and debatable, and the extent to which a special constitutional limitation goes is affected by the truth in respect to that fact, a widespread and long continued belief concerning it is worthy of consideration. We take judicial cognizance of all matters of general knowledge.

Coppage v. Kansas

Striking down legislation that outlawed so-called "yellow dog" contracts, the Court said:

An interference with this liberty so serious as that now under consideration, and so disturbing of equality of right, must be deemed to be arbitrary, unless it be supportable as a reasonable exercise of the police power of the State. But, notwithstanding the strong general presumption in favor of the validity of state laws, we do not think the statute in question, as construed and applied in this case, can be sustained as a legitimate exercise of that power. . . .

As to the interest of the employed, it is said by the Kansas Supreme Court (87 Kansas Rep., p. 759) to be a matter of common knowledge that "employees, as a rule, are not financially able to be as independent in making contracts for the sale of their labor as are employers in making contracts of purchase thereof." No doubt, wherever the right of private property exists, there must and will be inequalities of fortune; and thus it naturally happens that parties negotiating about a contract are not equally unhampered by circumstances. This applies to all contracts and not merely to that between employer and employee. Indeed a little reflection will show that wherever the right of private property and the right of free contract coexist each party when contracting is inevitably more or less influenced by the question whether he has much property, or little, or none; for the contract is made to the very end that each may gain something that he needs or desires more urgently than that which he proposes to give in exchange. And, since it is self-evident that, unless all things are held in common, some persons must have more property than others, it is from the nature of things impossible to uphold freedom of contract and the right of private property without at the same time recognizing as legitimate those inequalities of

fortune that are the necessary result of the exercise of those rights. But the Fourteenth Amendment, in declaring that a State shall not "deprive any person of life, liberty or property without due process of law," gives to each of these an equal sanction; it recognizes "liberty" and "property" as coexistent human rights, and debars the States from any unwarranted interference with either.

And since a State may not strike them down directly it is clear that it may not do so indirectly, as by declaring in effect that the public good requires the removal of those inequalities that are but the normal and inevitable result of their exercise, and then invoking the police power in order to remove the inequalities, without other object in view. The police power is broad, and not easily defined, but it cannot be given the wide scope that is here asserted for it, without in effect nullifying the constitutional guaranty.

West Coast Hotel Co. v. Parrish

Mr. Chief Justice Hughes delivered the opinion of the Court: This case presents the question of the constitutional validity of the minimum wage law of the State of Washington.

The Act . . . authorizes the fixing of minimum wages for women and minors. . . .

The appellant conducts a hotel. The appellee Elsie Parrish was employed as a chambermaid and (with her husband) brought this suit to recover the difference between the wages paid her and the minimum wage fixed pursuant to the state law . . . The Supreme Court of the State . . . sustained the statute. . . .

[The violation of the due process clause alleged by the appellant] is deprivation of freedom of contract. What is this freedom? The Constitution does not speak of freedom of contract. It speaks of liberty and prohibits the deprivation of liberty without due process of law. . . . But the liberty safeguarded is liberty in a social organization which requires the protection of law against the evils which menace the health, safety, morals and welfare of the people. Liberty under the Constitution is thus necessarily subject to the restraints of due process, and regulation which is reasonable in relation to its subject and is adopted in the interests of the community is due process.

This essential limitation of liberty in general governs freedom of contract in particular. . . .

We think . . . that the decision in the *Adkins* case [which in 1923 struck down congressional legislation setting minimum wages for women and children in the District of Columbia—Eds.] was a departure from the true application of the principles governing the regulation by the state of the relation of employer and employed. . . .

With full recognition of the earnestness and vigor which characterize the prevailing opinion in the *Adkins* case we find it impossible to reconcile that ruling with these well-considered declarations. What can be closer to the public interest than the health of women and their protection from unscrupulous and overreaching employers? And if the protection of women is a legitimate end of the exercise of state power, how can it be said that the requirement of the payment of a minimum wage fairly fixed in order to meet the very necessities of existence is not an admissible means to that end? . . .

There is an additional and compelling consideration which recent economic experience has brought into a strong light. The exploitation of a class of workers who are in an unequal position with respect to bargaining power and are thus relatively defenseless against the denial of a living wage is not only detrimental to their health and well being, but casts a direct burden for their support upon the community. What these workers lose in wages the taxpayers are called upon to pay. The bare cost of living must be met. We may take judicial notice of the unparalleled demands for relief which arose during the recent period of depression and still continue to an alarming extent despite the degree of economic recovery which has been achieved. It is unnecessary to cite official statistics to establish what is of common knowledge

through the length and breadth of the land. While in the instant case no factual brief has been presented, there is no reason to doubt that the state of Washington has encountered the same social problem that is present elsewhere. . . .

Our conclusion is that the case of *Adkins v. Children's Hospital* . . . should be, and it is, overruled.

JUSTICES SUTHERLAND, VAN DEVANTER, MCREYNOLDS and BUTLER dissented.

REVIEW AND DISCUSSION QUESTIONS

1. In the *Lochner* case, the Court states that it is sometimes permissible for government to interfere with employment contracts. Why did the Court not permit New York to do so in this case?
2. On what grounds does Holmes dissent from the majority's ruling in *Lochner*? Whose reasoning—Holmes's or the majority's—do you find more persuasive, and why?
3. Is the decision in *Muller v. Oregon* progressive because it protects employees from exploitation, or is it regressive because it holds that women deserve special protection not afforded to men?
4. In *Coppage v. Kansas*, the Court acknowledges that because employees have fewer material resources than do their employers, they are in a weaker bargaining position. Given this fact, one might question how free the contract between employee and employer really is. Nevertheless, the Court struck down legislation preventing employers from requiring, as a term of employment, that employees not belong to a labor union. What was the Court's reasoning?
5. What is the Supreme Court's position in *West Coast Hotel v. Parrish*, and how does it contrast with the positions upheld in the previous cases? On what grounds does it overrule the *Adkins* case?
6. Assess these cases from the points of view of Cohen and of Fried.

PART SIX

Philosophical Issues in Constitutional Law

CHAPTER 19

Church and State

Protecting basic rights was important to the framers of the Constitution. But what are the justifications and the limits of individual rights? One view, familiar to the framers, is that of the social contract, expressed here by John Locke in his *Letter Concerning Toleration.* In it, Locke argues that there is a sharp boundary between matters of legitimate concern to the governmental "magistrate" and matters of conscience. We then include three cases involving issues of religious freedom: the constitutionality of school prayer, the education of Amish children, and the teaching of "secular humanism" in public schools.

A Letter Concerning Toleration

John Locke

Few, if any, philosophers had as much influence on the thinking of the American framers as John Locke. Although today philosophers and political theorists focus more on his *Second Treatise of Government, A Letter Concerning Toleration* (1685) was much more influential at the time. In his *Letter,* Locke begins with a brief indication of his general theory of the powers of government. Legitimate government, he thinks, derives its powers from the consent of the governed, but such consent can best be understood by first imagining people in a pre-political "state of nature." There, Locke thinks, people would give up their natural right to life and property in order to gain greater protection of those rights by the state. The powers of the state, however, are limited by the original reasons that would motivate people to join and accept its authority. With this as background, Locke considers freedom of religion and conscience, an issue of central importance to the American framers as well. No government, he argues, should exercise force to promote or prohibit religious beliefs and practices.

Honored sir, . . . I esteem it above all things necessary to distinguish exactly the business of civil government from that of religion, and to settle the just bound that lie between the one and the other. If this be not done, there can be no end put to the controversies that will be always arising between those that have, or at least pretend to have, on the one side, a concernment for the

interest of men's souls, and, on the other side, a care of the commonwealth.

The commonwealth seems to me to be a society of men constituted only for the procuring, preserving, and advancing their own civil interests.

Civil interests I call life, liberty, health, and indolency of body; and the possession of outward things, such as money, lands, houses, furniture, and the like.

It is the duty of the civil magistrate, by the impartial execution of equal laws, to secure unto all the people in general, and to every one of his subjects in particular, the just possession of these things belonging to this life. If any one presume to violate the laws of public justice and equity, established for the preservation of those things, his presumption is to be checked by the fear of punishment, consisting of the deprivation or diminution of those civil interests, or goods, which otherwise he might and ought to enjoy. But seeing no man does willingly suffer himself to be punished by the deprivation of any part of his goods, and much less of his liberty or life, therefore is the magistrate armed with the force and strength of all his subjects, in order to the punishment of those that violate any other man's rights.

Now that the whole jurisdiction of the magistrate reaches only to these civil concernments, and that all civil power, right, and dominion, is bounded and confined to the only care of promoting these things; and that it neither can nor ought in any manner to be extended to the salvation of souls, these following considerations seem unto me abundantly to demonstrate.

First, because the care of souls is not committed to the civil magistrate, any more than to other men. It is not committed unto him, I say, by God: because it appears not that God has ever given any such authority to one man over another, as to compel any one to his religion. Nor can any such power be vested in the magistrate by the consent of the people, because no man can so far abandon the care of his own salvation as blindly to leave to the choice of any other, whether prince or subject, to prescribe to him what faith or worship he shall embrace. For no man can, if he would, conform his faith to the dictates of another. All the life and power of true religion consist in the inward and full persuasion of the mind; and faith is not faith without believing. Whatever profession we make, to whatever outward worship we conform, if we are not fully satisfied in our own mind that the one is true, and the other well pleasing unto God, such profession and such practice, far from being any furtherance, are indeed great obstacles to our salvation. For in this manner, instead of expiating other sins by the exercise of religion, I say, in offering thus unto God Almighty such a worship as we esteem to be displeasing unto him, we add unto the number of our other sins those also of hypocrisy, and contempt of his Divine Majesty.

In the second place, the care of souls cannot belong to the civil magistrate, because his power consists only in outward force; but true and saving religion consists in the inward persuasion of the mind, without which nothing can be acceptable to God. And such is the nature of the understanding, that it cannot be compelled to the belief of anything by outward force. Confiscation of estate, imprisonment, torments, nothing of that nature can have any such efficacy as to make men change the inward judgment that they have framed of things.

It may indeed be alleged that the magistrate may make use of arguments, and thereby draw the heterodox into the way of truth, and procure their salvation. I grant it; but this is common to him with other men. In teaching, instructing, and redressing the erroneous by reason, he may certainly do what becomes any good man to do. Magistracy does not oblige him to put off either humanity or Christianity; but it is one thing to persuade, another to command; one thing to press with arguments, another with penalties. This civil power alone has a right to do; to the other good-will is authority enough. Every man has commission to admonish, exhort, convince another of error, and, by reasoning, to draw him into truth; but to give laws, receive obedience, and compel with the sword, belongs to none but the magistrate. And upon this ground, I affirm that the magistrate's power extends not to the establishing of any articles of faith, or forms of worship, by the force of his laws. For laws are of no force at all without penalties, and penalties in this case are absolutely impertinent, because they are

not proper to convince the mind. Neither the profession of any articles of faith, nor the conformity to any outward form of worship (as has been already said), can be available to the salvation of souls, unless the truth of the one, and the acceptableness of the other unto God, be thoroughly believed by those that so profess and practise. But penalties are no way capable to produce such belief. It is only light and evidence that can work a change in men's opinions; which light can in no manner proceed from corporal sufferings, or any other outward penalties.

In the third place, the care of the salvation of men's souls cannot belong to the magistrate; because, though the rigour of laws and the force of penalties were capable to convince and change men's minds, yet would not that help at all to the salvation of their souls. For there being but one truth, one way to heaven, what hope is there that more men would be led into it if they had no rule but the religion of the court, and were put under the necessity to quit the light of their own reason, and oppose the dictates of their own consciences, and blindly to resign themselves up to the will of their governors, and to the religion which either ignorance, ambition, or superstition had chanced to establish in the countries where they were born? In the variety and contradiction of opinions in religion, wherein the princes of the world are as much divided as in their secular interests, the narrow way would be much straitened; one country alone would be in the right, and all the rest of the world put under an obligation of following their princes in the ways that lead to destruction; and that which heightens the absurdity, and very ill suits the notion of a Deity, men would owe their eternal happiness or misery to the places of their nativity.

These considerations, to omit many others that might have been urged to the same purpose, seem unto me sufficient to conclude that all the power of civil government relates only to men's civil interests, is confined to the care of the things of this world, and hath nothing to do with the world to come.

Let us now consider what a church is. A church, then, I take to be a voluntary society of men, joining themselves together of their own accord in order to the public worshipping of God in such manner as they judge acceptable to him, and effectual to the salvation of their souls.

I say it is a free and voluntary society. Nobody is born a member of any church; otherwise the religion of parents would descend unto children by the same right of inheritance as their temporal estates, and every one would hold his faith by the same tenure he does his lands, than which nothing can be imagined more absurd. Thus, therefore, that matter stands. No man by nature is bound unto any particular church or sect, but every one joins himself voluntarily to that society in which he believes he has found that profession and worship which is truly acceptable to God. The hope of salvation, as it was the only cause of his entrance into that communion, so it can be the only reason of his stay there. For if afterwards he discover anything either erroneous in the doctrine or incongruous in the worship of that society to which he has joined himself, why should it not be as free for him to go out as it was to enter? No member of a religious society can be tied with any other bonds but what proceed from the certain expectation of eternal life. A church, then, is a society of members voluntarily uniting to that end.

It follows now that we consider what is the power of this church, and unto what laws it is subject.

Forasmuch as no society, how free soever, or upon whatsoever slight occasion instituted, whether of philosophers for learning, of merchants for commerce, or of men of leisure for mutual conversation and discourse, no church or company, I say, can in the least subsist and hold together, but will presently dissolve and break in pieces, unless it be regulated by some laws, and the members all consent to observe some order. Place and time of meeting must be agreed on; rules for admitting and excluding members must be established; distinction of officers, and putting things into a regular course, and such-like, cannot be omitted. But since the joining together of several members into this church-society, as has already been demonstrated, is absolutely free and spontaneous, it necessarily follows that the right of making its laws can belong to none but the society itself; or, at least (which is the same thing), to

those whom the society by common consent has authorised thereunto. . . .

The end of a religious society (as has already been said) is the public worship of God, and, by means thereof, the acquisition of eternal life. All discipline ought therefore to tend to that end, and all ecclesiastical laws to be thereunto confined. Nothing ought nor can be transacted in this society relating to the possession of civil and worldly goods. No force is here to be made use of upon any occasion whatsoever. For force belongs wholly to the civil magistrate, and the possession of all outward goods is subject to his jurisdiction.

But, it may be asked, by what means then shall ecclesiastical laws be established, if they must be thus destitute of all compulsive power? I answer: They must be established by means suitable to the nature of such things, whereof the external profession and observation—if not proceeding from a thorough conviction and approbation of the mind—is altogether useless and unprofitable. The arms by which the members of this society are to be kept within their duty are exhortations, admonitions, and advices. If by these means the offenders will not be reclaimed, and the erroneous convinced, there remains nothing further to be done but that such stubborn and obstinate persons, who give no ground to hope for their reformation, should be cast out and separated from the society. This is the last and utmost force of ecclesiastical authority. No other punishment can thereby be inflicted than that, the relation ceasing between the body and the member which is cut off. The person so condemned ceases to be a part of that church.

These things being thus determined, let us inquire, in the next place: How far the duty of toleration extends, and what is required from every one by it?

And, first, I hold that no church is bound, by the duty of toleration, to retain any such person in her bosom as, after admonition, continues obstinately to offend against the laws of the society. For these being the condition of communion and the bond of the society, if the breach of them were permitted without any animadversion the society would immediately be thereby dissolved. . . .

Secondly, no private person has any right in any manner to prejudice another person in his civil enjoyments because he is of another church or religion. All the rights and franchises that belong to him as a man, or as a denizen, are inviolably to be preserved to him. These are not the business of religion. No violence nor injury is to be offered him, whether he be Christian or Pagan. Nay, we must not content ourselves with the narrow measures of bare justice; charity, bounty, and liberality must be added to it. This the Gospel enjoins, this reason directs, and this that natural fellowship we are born into requires of us. If any man err from the right way, it is his own misfortune, no injury to thee; nor therefore art thou to punish him in the things of this life because thou supposest he will be miserable in that which is to come.

What I say concerning the mutual toleration of private persons differing from one another in religion, I understand also of particular churches which stand, as it were, in the same relation to each other as private persons among themselves: nor has any one of them any manner of jurisdiction over any other; no, not even when the civil magistrate (as it sometimes happens) comes to be of this or the other communion. For the civil government can give no new right to the church, nor the church to the civil government. So that whether the magistrate join himself to any church, or separate from it, the church remains always as it was before—a free and voluntary society. It neither requires the power of the sword by the magistrate's coming to it, nor does it lose the right of instruction and excommunication by his going from it. This is the fundamental and immutable right of a spontaneous society—that it has power to remove any of its members who transgress the rules of its institution; but it cannot, by the accession of any new members, acquire any right of jurisdiction over those that are not joined with it. And therefore peace, equity, and friendship are always mutually to be observed by particular churches, in the same manner as by private persons, without any pretence of superiority or jurisdiction over one another.

. . . No peace and security, no, not so much as common friendship, can ever be established or preserved amongst men so long as this opinion prevails, that dominion is founded in grace and that religion is to be propagated by force of arms. . . .

REVIEW AND DISCUSSION QUESTIONS

1. What is the justification of the use of governmental power, according to Locke?
2. What are the three reasons why Locke thinks that the state must leave religious matters to the individual?
3. How does Locke understand religion, and how is that relevant to his discussion?
4. Do you agree with Locke's claim that "nobody is born a member of a church"? What do you think he means by it?

School Prayer

Engel v. Vitale

The Constitution's First Amendment has two separate religion clauses: one protecting the right of citizens to free exercise of their religion, the other preventing government from establishing a religion. Both clauses have led courts into difficult, controversial issues. Indeed, even today, nearly a half-century after school prayer was held unconstitutional, some still argue that this case was wrongly decided.

Black, J.: . . . The respondent Board of Education of Union Free School District No. 9, New Hyde Park, New York, acting in its official capacity under state law, directed the School District's principal to cause the following prayer to be said aloud by each class in the presence of a teacher at the beginning of each school day:

> Almighty God, we acknowledge our dependence upon Thee, and we beg Thy blessings upon us, our parents, our teachers and our Country.

This daily procedure was adopted on the recommendation of the State Board of Regents, a governmental agency created by the State Constitution to which the New York Legislature has granted broad supervisory, executive, and legislative powers over the State's public school system. These state officials composed the prayer which they recommended and published as a part of their "Statement on Moral and Spiritual Training in the Schools," saying: "We believe that this Statement will be subscribed to by all men and women of good will, and we call upon all of them to aid in giving life to our program."

Shortly after the practice of reciting the Regent's prayer was adopted by the School District, the parents of ten pupils brought this action in a New York State Court insisting that use of this official prayer in the public schools was contrary to the beliefs, religions, or religious practices of both themselves and their children. Among other things, these parents challenged the constitutionality of both the state law authorizing the School District to direct use of prayer in public schools and the School District's regulation ordering the recitation of this particular prayer on the ground that these actions of official governmental agencies violate . . . the First Amendment of the Federal Constitution . . .—a command which was "made applicable to the State of New York by the Fourteenth Amendment of the said Constitution." . . .

The petitioners contend among other things that the state laws requiring or permitting use of the Regents' prayer must be struck down as a violation of the Establishment Clause because that prayer was composed by governmental officials as a part of a governmental program to further religious beliefs. For this reason, petitioners argue, the State's use of the Regents' prayer in its public

370 U.S. 421 (1962).

school system breaches the constitutional wall of separation between Church and State. We agree with the contention since we think that the constitutional prohibition against laws respecting an establishment of religion must at least mean that in this country it is no part of the business of government to compose official prayers for any group of the American people to recite as a part of a religious program carried on by government.

It is a matter of history that this very practice of establishing governmentally composed prayers for religious services was one of the reasons which caused many of our early colonists to leave England and seek religious freedom in America. The Book of Common Prayer, which was created under governmental direction and which was approved by Acts of Parliament in 1548 and 1549, set out in minute detail the accepted form and content of prayer and other religious ceremonies to be used in the established, tax-supported Church of England. The controversies over the Book and what should be its content repeatedly threatened to disrupt the peace of that country as the accepted forms of prayer in the established church changed with the views of the particular ruler that happened to be in control at the time. Powerful groups representing some of the varying religious views of the people struggled among themselves to impress their particular views upon the Government and obtain amendments of the Book more suitable to their respective notions of how religious services should be conducted in order that the official religious establishment would advance their particular religious beliefs. Other groups, lacking the necessary political power to influence the government on the matter, decided to leave England and its established church and seek freedom in America from England's governmentally ordained and supported religion.

It is an unfortunate fact of history that when some of the very groups which had most strenuously opposed the established church of England found themselves sufficiently in control of colonial governments in this country to write their own prayers into law, they passed laws making their own religion the official religion of their respective colonies. Indeed, as late as the time of the Revolutionary War, there were established churches in at least eight of the thirteen former colonies and established religions in at least four of the other five. But the successful Revolution against the English political domination was shortly followed by intense opposition to the practice of establishing religion by law. This opposition crystallized rapidly into an effective political force in Virginia where the minority religious groups such as Presbyterians, Lutherans, Quakers and Baptists had gained such strength that the adherents to the established Episcopal Church were actually a minority themselves. In 1785–1786, those opposed to the established Church, led by James Madison and Thomas Jefferson, who, though themselves not members of any of these dissenting religious groups, opposed all religious establishments by law on grounds of principle, obtained the enactment of the famous "Virginia Bill for Religious Liberty" by which all religious groups were placed on an equal footing so far as the State was concerned. Similar though less far-reaching legislation was being considered and passed in other States.

By the time of the adoption of our Constitution, our history shows that there was a widespread awareness among many Americans of the dangers of a union of Church and State. These people knew, some of them from bitter personal experience, that one of the greatest dangers to the freedom of the individual to worship in his own way lay in the Government's placing its official stamp of approval upon one particular kind of prayer or one particular form of religious services. They knew the anguish, hardship and bitter strife that could come when zealous religious groups struggled with one another to obtain the Government's stamp of approval from each King, Queen, or Protector that came to temporary power. The Constitution was intended to avert a part of this danger by leaving the government of this country in the hands of the people rather than the hands of any monarch. But this safeguard was not enough. Our Founders were no more willing to let the content of their prayers and their privilege of praying whenever they pleased be influenced by the ballot box than they were to let these vital matters of personal conscience depend upon the succession of monarchs. The First Amendment was added to the Constitution to stand as a guarantee that neither the power nor the prestige of the Federal Government would be used

to control, support or influence the kinds of prayer the American people can say—that the people's religions must not be subjected to the pressures of government for change each time a new political administration is elected to office. Under that amendment's prohibition against governmental establishment of religion, as reinforced by the provisions of the Fourteenth Amendment, government in this country, be it state or federal, is without power to prescribe by law any particular form of prayer which is to be used as an official prayer in carrying on any program of governmentally sponsored religious activity.

There can be no doubt that New York's State prayer program officially establishes the religious beliefs embodied in the Regents' prayer. The respondents' argument to the contrary, which is largely based upon the contention that the Regents' prayer is "non-denominational" and the fact that the program, as modified and approved by state courts, does not require all pupils to recite the prayer but permits those who wish to do so to remain silent or be excused from the room, ignores the essential nature of the program's constitutional defects. Neither the fact that the prayer may be denominationally neutral nor the fact that its observance on the part of students is voluntary can serve to free it from the limitations of the Establishment Clause, as it might from the Free Exercise Clause, of the First Amendment, both of which are operative against the States by virtue of the Fourteenth Amendment. Although these two clauses may in certain instances overlap, they forbid two quite different kinds of governmental encroachment upon religious freedom. The Establishment Clause, unlike the Free Exercise Clause, does not depend upon any showing of direct governmental compulsion and is violated by the enactment of laws which establish an official religion whether those laws operate directly to coerce nonobserving individuals or not. This is not to say, of course, that laws officially prescribing a particular form of religious worship do not involve coercion of such individuals. When the power, prestige, and financial support of government is placed behind a particular religious belief, the indirect coercive pressure upon religious minorities to conform to the prevailing officially approved religion is plain. But the purposes underlying the Establishment Clause go much further than that. Its first and most immediate purpose rested on the belief that a union of government and religion tends to destroy government and to degrade religion. The history of governmentally established religion, both in England and in this country, showed that whenever government had allied itself with one particular religion, the inevitable result had been it had incurred the hatred, disrespect and even contempt of those who held contrary beliefs. That same history showed that many people had lost their respect for any religion that had relied upon the support of government to spread its faith. The establishment Clause thus stands as an expression of principle on the part of the Founders of our Constitution that religion is too personal, too sacred, too holy, to permit its "unhallowed perversion" by a civil magistrate. Another purpose of the Establishment Clause rested upon an awareness of the historical fact that governmentally established religions and religious persecutions go hand in hand. . . . It was in large part to get completely away from this sort of systematic religious persecution that the Founders brought into being our Nation, our Constitution, and our Bill of Rights with its prohibition against any governmental establishment of religion. The New York laws officially prescribing the Regents' prayer are inconsistent both with the purposes of the Establishment Clause and with the Establishment Clause itself.

It has been argued that to apply the Constitution in such a way as to prohibit state laws respecting an establishment of religious services in public schools is to indicate a hostility toward religion or toward prayer. Nothing, of course, could be more wrong. . . . It is neither sacrilegious nor antireligious to say that each separate government in this country should stay out of the business of writing or sanctioning official prayers and leave that purely religious function to the people themselves and to those the people choose to look for religious guidance.

. . . To those who may subscribe to the view that because the Regents' official prayer is so brief and general there can be no danger to religious freedom in its governmental establishment, however, it may be appropriate to say in the

words of James Madison, the author of the First Amendment:

> [I]t is proper to take alarm at the first experiment on our liberties. . . . Who does not see that the same authority which can establish Christianity, in exclusion of all other religions, may establish with the same ease any particular sect of Christians, in exclusion of all other Sects? That the same authority which can force a citizen to contribute three pence only of his property for the support of any one establishment, may force him to conform to any other establishment in all cases whatsoever?

The judgment of the Court of Appeals of New York is reversed and the cause remanded for further proceedings not inconsistent with this opinion.

Stewart, J. (dissenting): . . . With all respect, I think that the Court has misapplied a great constitutional principle. I cannot see how an "official religion" is established by letting those who want to say a prayer say it . . .

The Court's historical review of the quarrels over the Book of Common Prayer in England throws no light for me on the issue before us. England had then and has now an established church. Equally unenlightening, I think, is the history of the early establishment and later rejection of an official church in our own States. For we deal here not with the establishment of a state church, which would, of course, be constitutionally impermissible, but with whether school children who want to begin their day by joining in prayer must be prohibited from doing so. Moreover, I think that the Court's task . . . is not responsibly aided by the uncritical invocation of metaphors like the "wall of separation," a phrase nowhere to be found in the Constitution. What is relevant to the issue here is not the history of the established church in sixteenth century England or eighteenth century America, but the history of the religious traditions of our people, reflected in countless practices of the institutions and officials of our government.

At the opening of each day's session of this Court we stand, while one of our officials invokes the protection of God. Since the days of John Marshall our Crier has said, "God save the United States and this Honorable Court." Both the Senate and House of Representatives open their daily sessions with prayer. Each of our Presidents, from George Washington to John F. Kennedy, has upon assuming this Office asked the protection and help of God.

The Court today says that the state and federal governments are without constitutional power to prescribe any particular form of words to be recited by any group of the American people on any subject touching religion. One of the stanzas of "The Star-Spangled-Banner," made our National Anthem by Act of Congress in 1931, contains these verses:

> Blest with victory and peace, may the heav'n rescued land
> Praise the Pow'r that hath made and preserved us a nation!
> Then conquer we must, when our cause it is just
> And this be our motto "In God is our Trust."

In 1954 Congress added a phrase to the Pledge of Allegiance to the Flag so that it now contains the words "one Nation *under God,* indivisible, with liberty and justice for all." In 1952 Congress enacted legislation calling upon the President each year to proclaim a National Day of Prayer. Since 1865 the words "in god we trust" have been impressed on our coins.

Countless similar examples could be listed, but there is no need to belabor the obvious. It was all summed up by this court just ten years ago in a single sentence: "We are a religious people whose institutions presuppose a Supreme Being." *Zorach v. Clausen* (1952).

I do not believe that this Court, or the Congress, or the President has by the actions and practices I have mentioned established an "official religion" in violation of the Constitution. And I do not believe the State of New York has done so in this case. What each has done has been to recognize and to follow the deeply entrenched and highly cherished spiritual traditions of our Nation—traditions which come down to us from those who almost two hundred years ago avowed their "firm Reliance on the Protection of divine Providence" when they proclaimed the freedom and independence of this brave new world.

REVIEW AND DISCUSSION QUESTIONS

1. What reasons does Justice Black give for holding the New York school prayer unconstitutional?
2. Why does Justice Stewart reject Justice Black's argument?
3. Is the historical record described by Justice Black relevant or irrelevant to the argument?
4. How do you think Locke would have written this opinion?

Religious Freedom and Public Education

Wisconsin v. Yoder

This well-known case arose in response to a Wisconsin law requiring all children to be sent to school until the age of sixteen. The Amish parents of two children, ages fourteen and fifteen, refused to comply, arguing that compulsory school attendance beyond eighth grade violated their constitutional right of "free exercise" of religion protected by the First Amendment to the U.S. Constitution. The case went all the way to the Supreme Court, and the Court overturned the law, upholding the right of the Amish parents to guide the religious future and education of their children. In his dissenting opinion, Justice Douglas discusses problems associated with allowing parents to impose their religious notions on children as well as the possible effects of the Court's ruling on the children's educational development.

Mr. Chief Justice Burger: On complaint of the school district administrator for the public schools, respondents [Mr. and Mrs. Yoder] were charged, tried, and convicted of violating the compulsory-attendance law in Green County Court and were fined the sum of $5 each. Respondents defended on the ground that the application of the compulsory-attendance law violated their rights under the First and Fourteenth Amendments. The trial testimony showed that respondents believed, in accordance with the tenets of Old Order Amish communities generally, that their children's attendance at high school, public or private, was contrary to the Amish religion and way of life. . . . The State stipulated that respondents' religious beliefs were sincere.

In support of their position, respondents presented as expert witnesses scholars on religion and education whose testimony is uncontradicted. They expressed their opinions on the relationship of the Amish belief concerning school attendance to the more general tenets of their religion, and described the impact that compulsory high school attendance could have on the continued survival of Amish communities as they exist in the United States today. . . .

Amish beliefs require members of the community to make their living by farming or closely related activities. Broadly speaking, the Old Order Amish religion pervades and determines the entire mode of life of its adherents. . . .

Amish objection to formal education beyond the eighth grade is firmly grounded in these central religious concepts. They object to the high school, and higher education generally, because the values they teach are in marked variance with Amish values and the Amish way of life; they view secondary school education as an impermissible exposure of their children to a "worldly" influence in conflict with their beliefs. The high school tends to emphasize intellectual and scientific accomplishments, self-distinction, competitiveness, worldly success, and social life with other students. Amish society emphasizes informal learning-through-doing; a life of "goodness," rather than a life of intellect; wisdom, rather than technical knowledge; community

406 U.S. 205 (1972).

welfare, rather than competition; and separation from, rather than integration with, contemporary worldly society.

Formal high school education takes [Amish children] away from their community, physically and emotionally, during the crucial and formative adolescent period of life. During this period, the children must acquire Amish attitudes favoring manual work and self-reliance and the specific skills needed to perform the adult role of an Amish farmer or housewife. They must learn to enjoy physical labor. . . . And, at this time in life, the Amish child must also grow in his faith and his relationship to the Amish community if he is to be prepared to accept the heavy obligations imposed by adult baptism. . . .

The Amish do not object to elementary education through the first eight grades as a general proposition because they agree that their children must have basic skills in the "three Rs" in order to read the Bible, to be good farmers and citizens, and to be able to deal with non-Amish people when necessary in the course of daily affairs. They view such a basic education as acceptable because it does not significantly expose their children to worldly values or interfere with their development in the Amish community during the crucial adolescent period. . . .

On the basis of such considerations, [an expert] testified that compulsory high school attendance could not only result in great psychological harm to Amish children, because of the conflicts it would produce, but would also, in his opinion, ultimately result in the destruction of the Old Order Amish church community as it exists in the United States today. . . .

In order for Wisconsin to compel school attendance beyond the eighth grade against a claim that such attendance interferes with the practice of a legitimate religious belief, it must appear either that the State does not deny the free exercise of religious belief by its requirement, or that there is a state interest of sufficient magnitude to override the interest claiming protection under the Free Exercise Clause. . . .

A way of life, however virtuous and admirable, may not be interposed as a barrier to reasonable state regulation of education if it is based on purely secular considerations; to have the protection of the Religion Clauses, the claims must be rooted in religious belief. Although a determination of what is a "religious" belief or practice entitled to constitutional protection may present a most delicate question, the very concept of ordered liberty precludes allowing every person to make his own standards on matters of conduct in which society as a whole has important interests. Thus, if the Amish asserted their claims because of their subjective evaluation and rejection of the contemporary secular values accepted by the majority, much as Thoreau rejected the social values of his time and isolated himself at Walden Pond, their claims would not rest on a religious basis. Thoreau's choice was philosophical and personal rather than religious, and such belief does not rise to the demands of the Religion Clauses.

Giving no weight to such secular considerations, however, we see that the record in this case abundantly supports the claim that the traditional way of life of the Amish is not merely a matter of personal preference, but one of deep religious conviction, shared by an organized group, and intimately related to daily living. That the Old Order Amish daily life and religious practice stem from their faith is shown by the fact that it is in response to their literal interpretation of the Biblical injunction from the Epistle of Paul to the Romans, "be not conformed to this world. . . ."

Their way of life in a church-oriented community, separated from the outside world and "worldly" influences, their attachment to nature and the soil, is a way inherently simple and uncomplicated, albeit difficult to preserve against the pressure to conform. Their rejection of telephones, automobiles, radios, and television, their mode of dress, of speech, their habits of manual work do indeed set them apart from much of contemporary society; these customs are both symbolic and practical. . . .

The State advances two primary arguments in support of its system of compulsory education. It notes, as Thomas Jefferson pointed out early in our history, that some degree of education is necessary to prepare citizens to participate effectively and intelligently in our open political system if we are to preserve freedom and independence. Further, education prepares individuals to be self-reliant and self-sufficient participants in society. We accept these propositions.

However, the evidence adduced by the Amish in this case is persuasively to the effect that an additional one or two years of formal high school for Amish children in place of their long-established program of informal vocational education would do little to serve those interests. . . . It is one thing to say that compulsory education for a year or two beyond the eighth grade may be necessary when its goal is the preparation of the child for life in modern society as the majority live, but it is quite another if the goal of education be viewed as the preparation of the child for life in the separated agrarian community that is the keystone of the Amish faith. . . .

Whatever their idiosyncrasies as seen by the majority, the Amish community has been a highly successful social unit within our society, even if apart from the conventional "mainstream." Its members are productive and very law-abiding members of society; they reject public welfare in any of its usual modern forms. . . .

This case involves the fundamental interest of parents, as contrasted with that of the State, to guide the religious future and education of their children. The history and culture of Western civilization reflect a strong tradition of parental concern for the nurture and upbringing of their children. This primary role of the parents in the upbringing of their children is now established beyond debate as an enduring American tradition. . . .

To be sure, the power of the parent, even when linked to a free exercise claim, may be subject to limitation if it appears that parental decisions will jeopardize the health or safety of the child, or have a potential for significant social burdens. But in this case, the Amish have introduced persuasive evidence undermining the arguments the State has advanced to support its claims in terms of the welfare of the child and society as a whole.

Mr. Justice Douglas, dissenting in part: The Court's analysis assumes that the only interests at stake in the case are those of the Amish parents on the one hand, and those of the State on the other. The difficulty with this approach is that, despite the Court's claim, the parents are seeking to vindicate not only their own free exercise claims, but also those of their high-school-age children. . . .

No analysis of religious-liberty claims can take place in a vacuum. If the parents in this case are allowed a religious exemption, the inevitable effect is to impose the parents' notions of religious duty upon their children. Where the child is mature enough to express potentially conflicting desires, it would be an invasion of the child's rights to permit such an imposition without canvassing his views. . . . As the child has no other effective forum, it is in this litigation that his rights should be considered. And, if an Amish child desires to attend high school, and is mature enough to have that desire respected, the State may well be able to override the parents' religiously motivated objections.

This issue has never been squarely presented before today. Our opinions are full of talk about the power of the parents over the child's education. . . . And we have in the past analyzed similar conflicts between parent and State with little regard for the views of the child. . . . Recent cases, however, have clearly held that the children themselves have constitutionally protectible interests.

These children are "persons" within the meaning of the Bill of Rights. . . . While the parents, absent dissent, normally speak for the entire family, the education of the child is a matter on which the child will often have decided views. He may want to be a pianist or an astronaut or an oceanographer. To do so he will have to break from the Amish tradition. . . .

If a parent keeps his child out of school beyond the grade school, then the child will be forever barred from entry into the new and amazing world of diversity that we have today. . . .

[In the cases in which the Court held anti-polygamy laws constitutional,] action which the Court deemed to be antisocial could be punished even though it was grounded on deeply held and sincere religious convictions. What we to today, at least in this respect, opens the way to give organized religion a broader base than it has ever enjoyed.

In another way, however, the Court retreats when in reference to Henry Thoreau it says his "choice was philosophical and personal rather than religious, and such belief does not rise to the demands of the Religion Clauses." That is contrary

to what we held in *United States v. Seeger,* where we were concerned with the meaning of the words "religious training and belief" in the Selective Service Act, which were the basis of many conscientious objector claims. We said: "Within that phrase would come all sincere religious beliefs which are based upon a power or being, or upon a faith, to which all else is subordinate or upon which all else is ultimately dependent. The test might be stated in these words: A sincere and meaningful belief which occupies in the life of its possessor a place parallel to that filled by the God of those admittedly qualifying for the exemption comes within the statutory definition. This construction avoids imputing to Congress an intent to classify different religious beliefs, exempting some and excluding others, and is in accord with the well-established congressional policy of equal treatment for those whose opposition to service is grounded in their religious tenets."

REVIEW AND DISCUSSION QUESTIONS

1. Describe the legal issue in this case.
2. Does Justice Burger seem to rely more on religious freedom or the rights of parents over children? Explain.
3. On what basis does Justice Douglas dissent?
4. Does Justice Burger think the Constitution should be "neutral" among different religions? Explain.
5. Children will eventually have the right to make decisions for themselves in many areas, independent of what government or others may think. Can parents sometimes damage their children's future right to self-determination? Explain, giving examples.

Secular Humanism and Religious Establishment

Smith v. Board of Education of Mobile

The First Amendment guarantees citizens the right to exercise their religion and prevents government from "establishing" an official religion. But if religion is defined broadly enough, then it is possible to argue that by teaching "secular humanism" schools are establishing it as a religion. U.S. District Court Judge Brevard Hand makes that argument in the following case, *Smith v. Board of Education of Mobile.* The impetus for the case came in 1982 when the Supreme Court overruled an earlier decision by Judge Hand in which he had argued that a one-minute period of silent prayer in schools was not a violation of the establishment clause. Having lost that argument, Judge Hand did not let the issue die, suggesting in his opinion that should he be overruled (as he later was), he would then be willing to consider arguments that the school board had established another religion—secular humanism—in violation of the establishment clause. A group of religious conservatives took up the judge's offer and filed this case, *Smith v. Board of Education of Mobile.* After the trial, which included many expert witnesses on both sides, Judge Hand wrote a lengthy opinion contending that secular humanism is a religion and that the Mobile school board "established" it by requiring students to read certain textbooks. Although this decision was later overturned as well, Judge Hand's opinion in *Smith* raises important and difficult issues concerning the extent to which school texts that have the effect of undermining the values and religious beliefs of a community can be required.

---

655 F. Supp. 939 (S.D. Ala. 1987).

I. A FIRST AMENDMENT DEFINITION OF RELIGION

The Supreme Court has never stated an absolute definition of religion under the first amendment. . . .

The application of these principles to the question of what constitutes a religion under the first amendment indicates that the state may not decide the question by reference to the validity of the beliefs or practices involved. . . . The state must instead look to factors common to all religious movements to decide how to distinguish those ideologies worthy of the protection of the religious clauses from those which must seek refuge under other constitutional provisions.

· Any definition of religion must not be limited, therefore, to traditional religions, but must encompass systems of belief that are equivalent to them for the believer. . . . The Supreme Court has focused on such factors as a person's "ultimate concern," organization and social structure, and on equivalency to belief in a Supreme Deity. All of these are evidence of the type of belief a person holds. But all religious beliefs may be classified by the questions they raise and issues they address. Some of these matters overlap with nonreligious governmental concerns. A religion, however, approaches them on the basis of certain fundamental assumptions with which governments are unconcerned. These assumptions may be grouped as about: (1) the existence of supernatural and/or transcendent reality; (2) the nature of man; (3) the ultimate end, or goal or purpose of man's existence, both individually and collectively; (4) the purpose and nature of the universe. . . .

Whenever a belief system deals with fundamental questions of the nature of reality and man's relationship to reality, it deals with essentially religious questions. A religion need not posit a belief *in* a deity, or a belief *in* supernatural existence. A religious person adheres to some position on whether supernatural and/or transcendent reality exists at all, and if so, how, and if not, why. A mere "comprehensive world-view" or "way of life" is not by itself enough to identify a belief system as religious. A world-view may be merely economic, or sociological, and a person might choose to follow a "way of life" that ignores ultimate issues addressed by religions. . . .

There are also a number of characteristics exhibited by most known religious groups to which courts can look when trying to determine if a set of theories or system of ideas is religious in nature. First would be the sincerity of the adherents' claims. . . . Another factor is group organization and hierarchal structure, which evidence the social characteristics of a movement, and show that the adherents sincerely follow a theory of human relationship. Literary manifestations of a movement may also be important, particularly if they take the form of an authoritative text. Ritual and worship also would be significant because they would be evidence of the religion's belief about supernatural or transcendent reality. . . .

II. HUMANISM A RELIGION?

In the present case, the plaintiffs contend that a particular belief system fits within the first amendment definition of religion. . . . All of the experts, and the class representatives, agreed that this belief system [humanism] is a religion which: "makes a statement about supernatural existence a central pillar of its logic; defines the nature of man; and sets forth a goal or purpose for individual and collective human existence; and defines the nature of the universe, and thereby delimits its purpose."

It purports to establish a closed definition of reality; not closed in that adherents know everything, but in that everything is knowable: can be recognized by human intellect aided only by the devices of that intellect's own creation or discovery. The most important belief of this religion is its denial of the transcendent and/or supernatural: there is no God, no creator, no divinity. By force of logic the universe is thus self-existing, completely physical and hence, essentially knowable. Man is the product of evolutionary, physical forces. He is purely biological and has no supernatural or transcendent spiritual component or quality. Man's individual purpose is to seek and obtain personal fulfillment by freely developing every talent and ability, especially his rational intellect, to the highest level. Man's collective purpose is to seek the good life by the increase of

every person's freedom and potential for personal development.

In addition, humanism, as a belief system, erects a moral code and identifies the source of morality. This source is claimed to exist in humans and the social relationships of humans. Again, there is no spiritual or supernatural origin for morals: man is merely physical, and morals, the rules governing his private and social conduct, are founded only on man's actions, situation and environment. In addition to a moral code, certain attitudes and conduct are proscribed since they interfere with personal freedom and fulfillment. In particular any belief in a deity or adherence to a religious system that is theistic in any way is discouraged.

Secular humanism, or humanism in the sense of a religious belief system (as opposed to humanism as just an interest in the humanities), has organizational characteristics. . . .

These organizations publish magazines, newsletters and other periodicals, including *Free Inquiry, The Humanist,* and *Progressive World.* The entire body of thought has three key documents that furnish the text upon which the belief system rests as on a platform: *Humanist Manifesto I, Humanist Manifesto II,* and the *Secular Humanist Declaration.* . . .

To say that science is only concerned with data collected by the five senses as enhanced by technological devices of man's creation is to define science's limits. These are the parameters within which scientists function. However, to claim that there is nothing real beyond observable data is to make an assumption based not on science, but on faith, faith that observable data is all that is real. A statement that there is no transcendent or supernatural reality is a *religious* statement. . . .

This Court is holding that the promotion and advancement of a religious system occurs when one faith-theory is taught to the exclusion of others and this is prohibited by the first amendment religion clauses.

For purposes of the first amendment, secular humanism is a religious belief system, entitled to the protections of, and subject to the prohibitions of, the religion clauses. It is not a mere scientific methodology that may be promoted and advanced in the public schools.

III. RELIGIOUS PROMOTION IN TEXTBOOKS?

. . . [T]he Supreme Court has declared that teaching religious tenets in such a way as to promote or encourage a religion violates the religion clauses. This prohibition is not implicated by mere coincidence of ideas with religious tenets. Rather, there must be systematic, whether explicit or implicit, promotion of a belief system as a whole. The facts showed that the State of Alabama has on its state textbook list certain volumes that are being used by school systems in this state, which engage in such promotion. . . .

The virtually unanimous conclusion of the numerous witnesses, both expert and lay, party and non-party, was that textbooks in the fields examined were poor from an educational perspective. Mere rotten and inadequate textbooks, however, have not yet been determined to violate any constitutional provision, much less the religion clauses. . . . Their expert opinion was that religion was so deliberately underemphasized and ignored that theistic religions were effectively discriminated against and made to seem irrelevant and unimportant within the context of American history. . . . The religious influence on the abolitionist, woman's suffrage, temperance, modern civil rights and peace movements is ignored or diminished to insignificance. The role of religion in the lives of immigrants and minorities, especially southern blacks, is rarely mentioned. After the Civil War, religion is given almost no play. . . .

In addition to omitting particular historical events with religious significance, these books uniformly ignore the religious aspect of most American culture. The vast majority of Americans, for most of our history, have lived in a society in which religion was a part of daily life. . . . For many people, religion is still this important. One would never know it by reading these books. Religion, where treated at all, is generally represented as a private matter, only influencing American public life at some extraordinary moments. This view of religion is one humanists have been seeking to instill for fifty years. These books assist that effort by perpetuating an inaccurate historical picture. . . .

According to humanistic psychology, as with humanism generally, man is the center of the

universe and all existence. Morals are a matter of taste, dependent upon whether the consequences of actions satisfy human "needs." These needs are always defined as purely temporal and non-supernatural. . . .

The [social studies] books do not state that this is a *theory* of the way humans make choices, they teach the student that things *are* this way. . . . The books teach that the student must determine right and wrong based only on his own experience, feelings and "values." These "values" are described as originating from within. A description of the origin of morals must be based on a faith assumption: a religious dogma. The books are not simply claiming that a moral rule must be internally accepted before it becomes meaningful, because this is true of *all* facts *and* beliefs. . . . The books repeat, over and over, that the decision is "yours alone," or is "purely personal" or that "only you can decide." The emphasis and overall approach implies, and would cause any reasonable, thinking student to infer, that the book is teaching that moral choices are just a matter of preference, because, as the books say, "you are the most important person in your life." . . . This faith assumes that self-actualization is the goal of every human being, that man has no supernatural attributes or component, that there are only temporal and physical consequences for man's actions, and that these results, alone, determine the morality of an action. This belief strikes at the heart of many theistic religions' beliefs that certain actions are in and of themselves immoral, *whatever the consequences,* and that, in addition, actions will have extratemporal consequences. The Court is not holding that high school . . . books must not discuss various theories of human psychology. But it must not present faith based systems to the exclusion of other faith based systems, it must not present one as true and the other as false, and it *must* use a comparative approach to withstand constitutional scrutiny. . . .

With these [history and social studies] books, the State of Alabama has overstepped its mark, and must withdraw to perform its proper non-religious functions. . . . The Court will enter an order to prohibit further use of the books.

REVIEW AND DISCUSSION QUESTIONS

1. Describe the legal issue that is before the court in this case.
2. How does Judge Hand understand the legal definition of *religion?*
3. Why does he think humanism is a religion?
4. What does Judge Hand find in the textbooks that leads him to believe that they "establish" a religion?
5. The court of appeals overturned Judge Hand's decision, and the Supreme Court agreed with the appeals court. Is Judge Hand's opinion consistent with the Supreme Court's ruling in *Wisconsin v. Yoder?* Explain.
6. In another, related, case the Supreme Court ruled that schools can teach evolutionary theory despite the fact that it is incompatible with some students' religious views. How would you decide such a case? Explain.

CHAPTER 20

Personal Liberty and Privacy

The freedom of the individual to choose and to act as he or she sees fit is a fundamental value, one that seems deeply rooted in our political and constitutional tradition. Yet any society must place limits on individual conduct. In his classic essay *On Liberty*, John Stuart Mill addresses the fundamental question, What are the proper limits to society's power over the individual? Mill's answer is that society may interfere with an individual's actions only to prevent the individual from harming others. Society may not compel the individual to do (or not do) something solely because society judges that it would be in the individual's own interest to do it. Closely related to Mill's liberty principle is the right to privacy—in particular, the right to make personal decisions autonomously, free from the interference of the state—and debates over the exact nature and extent of a constitutional right to privacy have loomed large in recent Supreme Court decisions, especially with regard to consensual sexual conduct and same-sex marriage.

On Liberty

John Stuart Mill

John Stuart Mill (1806–73) had an unusual life by almost any account. His father, James Mill, who was friends with economist David Ricardo and legal theorist John Austin, was among the most devoted followers of Jeremy Bentham, the utilitarian philosopher. Like Bentham, the senior Mill held that human behavior is best understood in terms of self-interest and that government serves the common good by promoting the happiness of its citizens.

James resolved to raise his son John Stuart according to sound utilitarian principles. With Bentham's help, he developed a plan that included a rigorous tutoring program and keeping the boy away from other children. Young John Stuart was a brilliant student. By the age of three, he had begun learning Greek. At eight, he learned Latin and pursued mathematics and history. By the age of twelve, he was studying logic and political economy, and at fifteen he studied law at University College, London, with John Austin. At twenty-four Mill began a lifelong friendship and intellectual collaboration with Harriet

Taylor. Although they were close companions, Harriet Taylor remained married to another man for two decades. When he died, she and John Stuart Mill married. After their marriage the couple withdrew from society and the gossip they had endured for years, living happily for seven years until Harriet's death. Mill served briefly as a member of Parliament and died in France (where he had bought a house near the cemetery in which Harriet was buried).

Mill's influence has been tremendous; he wrote important books in logic, philosophy of science, and economics as well as ethics and political philosophy. One of his works, *Utilitarianism,* remains a classic statement of that important theory. None of his books, however, has had a greater impact than *On Liberty.* In fact, *On Liberty* should be thought of as a joint work authored by both Mill and Harriet Taylor. Mill had tremendous respect for her intellectual ability and wrote in a letter that there is not a sentence in the book that the two did not go over together many times. Harriet Taylor also had a major role in some of his other works, including *On the Subjection of Women.*

In this selection from *On Liberty*, Mill examines a fundamental question of political philosophy: What are the limits of society's power over the individual? In answering this question, Mill defends, he says, "one very simple principle": namely, that society is justified in interfering with the liberty of the individual only on grounds of self-protection, that is, only if the individual's actions threaten harm to others.

INTRODUCTORY

The subject of this essay is . . . civil, or social liberty: the nature and limits of power which can be legitimately exercised by society over the individual.

. . . [In] old times this contest was between subjects, or some classes of subjects, and the government. By liberty, was meant protection against the tyranny of the political rulers. The rulers were conceived (except in some of the popular governments of Greece) as in a necessarily antagonistic position to the people whom they ruled. . . .

A time, however, came, in the progress of human affairs, when men ceased to think it a necessity of nature that their governors should be an independent power, opposed in interest to themselves. It appeared to them much better that the various magistrates of the State should be their tenants or delegates, revocable at their pleasure. In that way alone, it seemed, could they have complete security that the powers of government would never be abused to their disadvantage. . . . What was now wanted was, that the rulers should be identified with the people; that their interest and will should be the interest and will of the nation. The nation did not need to be protected against its own will. . . .

But in political and philosophical theories, as well as in persons, success discloses faults and infirmities which failure might have concealed from observation. . . . It was now perceived that such phrases as "self-government," and "the power of the people over themselves," do not express the true state of the case. The "people" who exercise the power are not always the same people with those over whom it is exercised; and the "self-government" spoken of is not the government of each by himself, but of each by all the rest. The will of the people, moreover, practically means the will of the most numerous or the most active *part* of the people—the majority, or those who succeed in making themselves accepted as the majority; the people, consequently, *may* desire to oppress a part of their number, and precautions are as much needed against this as against any other abuse of power. The limitation, therefore, of the power of government over individuals loses none of its importance when the holders of power are regularly accountable to the community, that is, to the strongest party therein. . . .

The object of this essay is to assert one very simple principle, as entitled to govern absolutely the dealings of society with the individual in the way of compulsion and control, whether the means used by physical force in the form of legal penalties or the moral coercion of public opinion. That principle is, that the sole end for which mankind are warranted, individually or collectively, in interfering with the liberty of action of any of their number, is self-protection. That the only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others. His own good, either physical

or moral, is not a sufficient warrant. He cannot rightfully be compelled to do or forbear because it will be better for him to do so, because it will make him happier, because, in the opinions of others, to do so would be wise or even right. These are good reasons for remonstrating with him, or reasoning with him, or persuading him, or entreating him, but not for compelling him or visiting him with any evil in case he do otherwise. To justify that, the conduct from which it is desired to deter him must be calculated to produce evil to someone else. The only part of the conduct of anyone, for which he is amenable to society, is that which concerns others. In the part which merely concerns himself, his independence is, of right, absolute. Over himself, over his own body and mind, the individual is sovereign.

It is, perhaps, hardly necessary to say that this doctrine is meant to apply only to human beings in the maturity of their faculties. We are not speaking of children, or of young persons below the age which the law may fix as that of manhood or womanhood. Those who are still in a state to require being taken care of by others, must be protected against their own actions as well as against external injury. For the same reason, we may leave out of consideration those backward states of society in which the race itself may be considered as in its nonage. The early difficulties in the way of spontaneous progress are so great, and there is seldom any choice of means for overcoming them; and a ruler full of the spirit of improvement is warranted in the use of any expedients that will attain an end, perhaps otherwise unattainable. Despotism is a legitimate mode of government in dealing with barbarians, provided the end be their improvement, and the means justified by actually effecting that end. Liberty, as a principle, has no application to any state of things anterior to the time when mankind have become capable of being improved by free and equal discussion. Until then, there is nothing for them but implicit obedience to an Akbar or a Charlemagne, if they are so fortunate as to find one. But as soon as mankind have attained the capacity of being guided to their own improvement by conviction or persuasion (a period long since reached in all nations with whom we need here concern ourselves), compulsion, either in the direct form or in that of pains and penalties for noncompliance, is no longer admissible as a means to their own good, and justifiable only for the security of others.

It is proper to state that I forego any advantage which could be derived to my argument from the idea of abstract right, as a thing independent of utility. I regard utility as the ultimate appeal on all ethical questions; but it must be utility in the largest sense, grounded on the permanent interests of a man as a progressive being. These interests, I contend, authorized the subjection of individual spontaneity to external control, only in respect to those actions of each which concern the interest of other people. If anyone does an act hurtful to others, there is a *prima facie* case for punishing him, by law, or, where legal penalties are not safely applicable, by general disapprobation. There are also many positive acts for the benefit of others, which he may rightfully be compelled to perform: such as to give evidence in a court of justice; to bear his fair share in the common defense, or in any other joint work necessary to the interest of the society of which he enjoys the protection; and to perform certain acts of individual beneficence, such as saving a fellow-creature's life, or interposing to protect the defenseless against ill-usage, things which whenever it is obviously a man's duty to do, he may rightfully be made responsible to society for not doing. A person may cause evil to others not only by his actions but by his inaction, and in either case he is justly accountable to them for the injury. The latter case, it is true, requires a much more cautious exercise of compulsion than the former. . . .

. . . This, then is the appropriate region of human liberty. It comprises, *first,* the inward domain of consciousness; demanding liberty of conscience in the most comprehensive sense; liberty of thought and feeling; absolute freedom of opinion and sentiment on all subjects, practical or speculative, scientific, moral or theological. The liberty of expressing and publishing opinions may seem to fall under a different principle, since it belongs to that part of the conduct of an individual which concerns other people; but, being almost of as much importance as the liberty of thought itself, and resting in great part on the same reasons, is practically inseparable from it. *Secondly,* the principle requires liberty of tastes and pursuits; of framing the plan of our life to suit

our own character; of doing as we like, subject to such consequences as may follow: without impediment from our fellow-creatures, so long as what we do does not harm them, even though they should think our conduct foolish, perverse, to wrong. *Thirdly,* from this liberty of each individual, follows the liberty, within the same limits, of combinations among individuals; freedom to unite, for any purpose not involving harm to others: the persons combining being supposed to be of full age, and not forced or deceived.

No society in which these liberties are not, on the whole, respected, is free, whatever may be its form of government; and none is completely free in which they do not exist absolute and unqualified. The only freedom which deserves the name, is that of pursuing our own good in our own way, as long as we do not attempt to deprive others of theirs, or impede their efforts to obtain it. Each is the proper guardian of his own health, whether bodily, or mental and spiritual. Mankind are greater gainers by suffering each other to live as seems good to themselves, than by compelling each to live as seems good to the rest. . . . [Mill now discusses freedom of speech, which is reprinted beginning on p. 589—Eds.]

OF THE LIMITS TO THE AUTHORITY OF SOCIETY OVER THE INDIVIDUAL

What, then, is the rightful limit to the sovereignty of the individual over himself? Where does the authority of society begin? How much of human life should be assigned to individuality, and how much to society?

Each will receive its proper share, if each has that which more particularly concerns it. To individuality should belong the part of life in which it is chiefly the individual that is interested; to society, the part which chiefly interests society.

Though society is not founded on a contract, and though no good purpose is answered by inventing a contract in order to deduce social obligations from it, everyone who receives the protection of society owes a return for the benefit, and the fact of living in society renders it indispensable that each should be bound to observe a certain line of conduct towards the rest. This conduct consists, *first,* in not injuring the interests of one another; or rather certain interests, which either by express legal provision or by tacit understanding, ought to be considered as rights; and *secondly,* in each person's bearing his share (to be fixed on some equitable principle) of the labors and sacrifices incurred for defending the society or its members from injury and molestation. These conditions society is justified in enforcing, at all costs to those who endeavor to withhold fulfillment. Nor is this all that society may do. The acts of an individual may be hurtful to others, or wanting in due consideration for their welfare, without going to the length of violating any of their constituted rights. The offender may then be justly punished by opinion, though not by law. As soon as any part of a person's conduct affects prejudicially the interests of others, society has jurisdiction over it, and the question whether the general welfare will or will not be promoted by interfering with it, becomes open to discussion. But there is no room for entertaining any such question when a person's conduct affects the interests of no persons besides himself, or need not affect them unless they like (all the persons concerned being of full age, and the ordinary amount of understanding). In all such cases, there should be perfect freedom, legal and social, to do the action and stand the consequences.

It would be a great misunderstanding of this doctrine to suppose that it is one of selfish indifference, which pretends that human beings have no business with each other's conduct in life, and that they should not concern themselves about the well-doing or well-being of one another, unless their own interest is involved. Instead of any diminution, there is need of a great increase of disinterested exertion to promote the good of others. But disinterested benevolence can find other instruments to persuade people to their good than whips and scourges, either of the literal or the metaphorical sort. I am the last person to undervalue the self-regarding virtues: they are only second in importance, if even second, to the social. It is equally the business of education to cultivate both. But even education works by conviction and persuasion as well as by compulsion, and it is by the former only that, when the period of education is passed, the self-regarding virtues should be inculcated. Human beings owe to each other help to distinguish the

better from the worse, and encouragement to choose the former and avoid the latter. They should be forever stimulating each other to increased exercise of their higher faculties, and increased direction of their feelings and aims towards wise instead of foolish, elevating instead of degrading, objects and contemplations. But neither one person, nor any number of persons, is warranted in saying to another human creature of ripe years, that he shall not do with his life for his own benefit what he chooses to do with it. He is the person most interested in his own well-being: the interest which any other person, except in cases of strong personal attachment, can have in it, is trifling, compared with that which he himself has; the interest which society has in him individually (except as to conduct to others) is fractional, and altogether indirect; while with respect to his own feelings and circumstances, the most ordinary man or woman has means of knowledge immeasurably surpassing those that can be possessed by anyone else. The interference of society to overrule his judgment and purposes in what only regards himself must be grounded on general presumptions; which may be altogether wrong, and even if right, are as likely as not to be misapplied to individual cases, by person no better acquainted with the circumstances of such cases than those who look at them merely from without. In this department, therefore, of human affairs, individuality has its proper field of action. In the conduct of human beings towards one another it is necessary that general rules should for the most part be observed, in order that people may know what they have to expect; but in each person's own concerns his individual spontaneity is entitled to free exercise. Considerations to aid his judgment, exhortations to strengthen his will, may be offered to him, even obtruded on him, by others: but he himself is the final judge. All errors which he is likely to commit against advice and warning are far outweighed by the evil of allowing others to constrain him to what they deem his good. . . .

Though doing no wrong to anyone, a person may so act as to compel us to judge him, and feel to him, as a fool, or as a being of an inferior order; and since this judgment and feeling are a fact which he would prefer to avoid, it is doing him a service to warn him of it beforehand, as of any other disagreeable consequence to which he exposes himself. . . . We have a right, also, in various ways, to act upon our unfavorable opinion of anyone, not to the oppression of his individuality, but in the exercise of ours. We are not bound, for example, to seek his society; we have a right to avoid it (though not to parade the avoidance), for we have a right to choose the society most acceptable to us. We have a right, and it may be our duty, to caution others against him, if we think his example or conversation likely to have a pernicious effect on those with whom he associates. We may give others a preference over him in optional good offices, except those which tend to his improvement. In these various modes a person may suffer very severe penalties at the hands of others for faults which directly concern only himself; but he suffers these penalties only in so far as they are the natural and, as it were, the spontaneous consequences of the faults themselves, not because they are purposely inflicted on him for the sake of punishment. . . .

What I contend for is, that the inconveniences which are strictly inseparable from the unfavorable judgment of others, are the only ones to which a person should ever be subjected for that portion of his conduct and character which concerns his own good, but which does not affect the interest of others in their relations with him. Acts injurious to others require a totally different treatment. Encroachment on their rights; infliction on them of any loss or damage not justified by his own rights; falsehood or duplicity in dealing with them; unfair or ungenerous use of advantages over them; even selfish abstinence from defending them against injury—these are fit objects of moral reprobation, and, in grave cases, of moral retribution and punishment. And not only these acts, but the dispositions which lead to them, are properly immoral, and fit subjects of disapprobation which may rise to abhorrence. . . .

The distinction here pointed out between the part of a person's life which concerns only himself, and that which concerns others, many persons will refuse to admit. How (it may be asked) can any part of the conduct of a member of society be a matter of indifference to the other members? No person is an entirely isolated being; it is impossible for a person to do anything seriously or permanently hurtful to himself, without mischief reaching at least to his near connections, and often far beyond them. If he

injures his property, he does harm to those who directly or indirectly derived support from it, and usually diminishes, by a greater or less amount, the general resources of the community. If he deteriorates his bodily or mental faculties, he not only brings evil upon all who depended on him for any portion of their happiness, but disqualifies himself for rendering the services which he owes to his fellow-creatures generally; perhaps becomes a burden on their affection or benevolence; and if such conduct were very frequent, hardly an offense that is committed would detract more from the general sum of good. Finally, if by his vices or follies a person does no direct harm to others, he is nevertheless (it may be said) injurious by his example; and ought to be compelled to control himself, for the sake of those whom the sight or knowledge of his conduct might corrupt or mislead.

And even (it will be added) if the consequences of misconduct could be confined to the vicious or thoughtless individual, ought society to abandon to their own guidance those who are manifestly unfit for it? If protection against themselves is confessedly due to children and persons under age, is not society equally bound to afford it to persons of mature years who are equally incapable of self-government? If gambling, or drunkenness, or incontinence, or idleness, or uncleanliness, are as injurious to happiness, and as great a hindrance to improvement, as many or most of the acts prohibited by law, why (it may be asked) should not law, so far as is consistent with practicability and social convenience, endeavor to repress these also? . . .

I fully admit that the mischief which a person does to himself may seriously affect, both through their sympathies and their interests, those nearly connected with him and, in a minor degree, society at large. When, by conduct of this sort, a person is led to violate a distinct and assignable obligation to any other person or persons, the case is taken out of the self-regarding class and becomes amenable to moral disapprobation in the proper sense of the term. If, for example, a man, through intemperance or extravagance, becomes unable to pay his debts, or, having undertaken the moral responsibility of a family, becomes from the same cause incapable of supporting or educating them, he is deservedly reprobated, and might be justly punished; but it is for the breach of duty to his family or creditors, not for the extravagance. . . . No person ought to be punished simply for being drunk; but a soldier or policeman should be punished for being drunk on duty. Whenever, in short, there is a definite damage, or a definite risk of damage, either to an individual or to the public, the case is taken out of the province of liberty and placed in that of morality or law.

But with regard to the merely contingent or, as it may be called, constructive injury which a person causes to society by conduct which neither violates any specific duty to the public, nor occasions perceptible hurt to any assignable individual except himself, the inconvenience is one which society can afford to bear, for the sake of the greater good of human freedom. If grown persons are to be punished for not taking proper care of themselves, I would rather it were for their own sake than under pretense of preventing them from impairing their capacity or rendering to society benefits which society does not pretend it has a right to exact. But I cannot consent to argue the point as if society had no means of bringing its weaker members up to its ordinary standard of rational conduct, except waiting till they do something irrational, and then punishing them, legally or morally, for it. Society has had absolute power over them during all the early portion of their existence; it has had the whole period of childhood and nonage in which to try whether it could make them capable of rational conduct in life. The existing generation is master both of the training and the entire circumstances of the generation to come; it cannot indeed make them perfectly wise and good, because it is itself so lamentably deficient in goodness and wisdom; and its best efforts are not always, in individual cases, its most successful ones; but it is perfectly well able to make the rising generation, as a whole, as good as, and a little better than, itself. If society lets any considerable number of its members grow up mere children, incapable of being acted on by rational consideration of distant motives, society has itself to blame for the consequences. Armed not only with all the powers of education, but with the ascendency which the authority of a received opinion always exercises over the minds who are least fitted to judge for themselves, and aided by the *natural* penalties which cannot be prevented

from falling on those who incur the distaste or the contempt of those who know them—let not society pretend that it needs, besides all this, the power to issue commands and enforce obedience in the personal concerns of individuals in which, on all principles of justice and policy, the decision ought to rest with those who are to abide the consequences. Nor is there anything which tends more to discredit and frustrate the better means of influencing conduct than a resort to the worse. If there be among those whom it is attempted to coerce into prudence or temperance any of the material of which vigorous and independent characters are made, they will infallibly rebel against the yoke. No such person will ever feel that others have a right to control him in his concerns, such as they have to prevent him from injuring them in theirs; and it easily comes to be considered a mark of spirit and courage to fly in the face of such usurped authority and do with ostentation the exact opposite of what it enjoins, as in the fashion of grossness which succeeded, in the time of Charles II, to the fanatical moral intolerance of the Puritans. With respect to what is said of the necessity of protecting society from the bad example set to others by the vicious or the self-indulgent, it is true that bad example may have a pernicious effect, especially the example of doing wrong to others with impunity to the wrongdoer. But we are now speaking of conduct which, while it does no wrong to others, is supposed to do great harm to the agent himself; and I do not see how those who believe this can think otherwise than that the example, on the whole, must be more salutary than hurtful since, if it displays the misconduct, it displays also the painful or degrading consequences which, if the conduct is justly censured, must be supposed to be in all or more cases attendant on it.

But the strongest of all the argument against the interference of the public with purely personal conduct is that, when it does interfere, the odds are that it interferes wrongly and in the wrong place. On questions of social morality, of duty to others, the opinion of the public, that is, of an overruling majority, though often wrong, is likely to be still oftener right, because on such questions they are only required to judge of their own interests, of the manner in which some mode of conduct, if allowed to be practiced, would affect themselves. But the opinion of a similar majority, imposed as a law on the minority, on questions of self-regarding conduct is quite as likely to be wrong as right, for in these cases public opinion means, at the best, some people's opinion of what is good or bad for other people, while very often it does not even mean that—the public, with the most perfect indifference, passing over the pleasure or convenience of those whose conduct they censure and considering only their own preference. There are many who consider as an injury to themselves any conduct which they have a distaste for, and resent it as an outrage to their feelings; as a religious bigot, when charged with disregarding the religious feelings of others, has been known to retort that they disregard his feelings by persisting in their abominable worship or creed. But there is no parity between the feeling of a person for his own opinion and the feeling of another who is offended at his holding it, no more than between the desire of a thief to take a purse and the desire of the right owner to keep it. And a person's taste is as much his own peculiar concern as his opinion or his purse.

REVIEW AND DISCUSSION QUESTIONS

1. Mill implicitly believes that the good life involves more than mere contentment. What else do you think it involves for him?
2. How does Mill think most people make decisions about how they will live their lives?
3. Mill admits many acts may affect others but insists that unless they also "harm" them government may not intervene. How does Mill understand "harm"?
4. Describe how Mill thinks society *may* try to influence people whose actions are harmless, yet either immoral or not in their true interest.
5. What reasons does Mill give to support his claim that harming others is necessary to justify punishment?

6. Sometimes, Mill says, society is wise not to interfere even though an act *is* harmful to another. Give an example of such an act, then explain why Mill thinks society should not interfere.
7. In some cases interfering with an individual's self-regarding conduct might increase society's overall happiness or well-being. If so, how can Mill square his almost absolute defense of liberty with his utilitarianism?

The Right to Privacy

Griswold v. Connecticut

Some years ago Connecticut law mandated fine or imprisonment of "any person who uses any drug, medicinal article or instrument for the purpose of preventing conception" and further stipulated that any person who "aid, abets, counsels, causes, hires or commands" another person to commit this offense was subject to prosecution and punishment "as if he were the principal offender." Griswold, the Executive Director of Planned Parenthood, and a professor of medicine named Buxton were found guilty of violating the statute and fined $100 apiece. Both appealed, arguing that the law violated their rights under the Fourteenth Amendment. The Supreme Court's decision in this case has provided the grounding for a number of other important decisions, including *Roe v. Wade,* which upheld a woman's right to abortion.

Mr. Justice Douglas delivered the opinion of the Court: We are met with a wide range of questions that implicate the Due Process Clause of the 14th Amendment. Overtones of some arguments suggest that *Lochner v. New York,* 198 U.S. 45, should be our guide. But we decline that invitation. We do not sit as a super-legislature to determine the wisdom, need, and propriety of laws that touch economic problems, business affairs, or social conditions. This law, however, operates directly on an intimate relation of husband and wife and their physician's role in one aspect of that relation.

The association of people is not mentioned in the Constitution nor in the Bill of Rights. The right to educate a child in a school of the parents' choice—whether public or private or parochial—is also not mentioned. Nor is the right to study any particular subject or any foreign language. Yet the First Amendment has been construed to include certain of those rights. The State may not, consistently with the spirit of the First Amendment, contract the spectrum of available knowledge. The right of freedom of speech and press includes not only the right to utter or to print, but the right to distribute, the right to receive, the right to read [and] freedom of inquiry, freedom of thought, and freedom to [teach]. Without those peripheral rights the specific rights would be less secure. In *NAACP v. Alabama,* we protected the "freedom to associate and privacy in one's association," noting that freedom of association was a peripheral First Amendment right. Disclosure of membership lists of a constitutionally valid association, we held, was invalid. In other words, the First Amendment has a penumbra where privacy is protected from governmental intrusion. In like context, we have protected forms of "association" that are not political in the customary sense but pertain to the social, legal, and economic benefit of the members. The right of association, while it is not expressly included in the First Amendment, is necessary in making the express guarantees fully meaningful.

The foregoing cases suggest that specific guarantees in the Bill of Rights have penumbras, formed by emanations from those guarantees

381 U.S. 479 (1965).

that help give them life and substance. Various guarantees create zones of privacy. The right of association contained in the penumbra of the First Amendment is one, as we have seen. The Third Amendment in its prohibition against the quartering of soldiers "in any house" in time of peace without the consent of the owner is another facet of that privacy. The Fourth Amendment explicitly affirms the "right of the people to be secure in their persons, houses, papers, and effects against unreasonable searches and seizures." The Fifth Amendment in its Self-Incrimination Clause enables the citizen to create a zone of privacy which government may not force him to surrender to his detriment. The Ninth Amendment provides: "The enumeration in the Constitution, of certain rights, shall not be construed to deny or disparage others retained by the people." The Fourth and Fifth Amendments protect against all governmental invasions "of the sanctity of a man's home and the privacies of life." We recently referred to the Fourth Amendment as creating a "right to privacy, no less important than any other right carefully and particularly reserved to the people." We have had many controversies over these penumbral rights of "privacy and repose." These cases bear witness that the right of privacy which presses for recognition here is a legitimate one.

The present case, then, concerns a relationship lying within the zone of privacy created by several fundamental constitutional guarantees. And it concerns a law which, in forbidding the *use* of contraceptives rather than regulating their manufacture or sale, seeks to achieve its goals by means having a maximum destructive impact upon that relationship. Such a law cannot stand in light of the familiar principle that a "governmental purpose to control or prevent activities constitutionally subject to state regulation may not be achieved by means which sweep unnecessarily broadly and thereby invade the area of protected freedoms." Would we allow the police to search the sacred precincts of marital bedrooms for telltale signs of the use of contraceptives? The very idea is repulsive to the notions of privacy surrounding the marriage relationship.

We deal with a right of privacy older than the Bill of Rights. Marriage is a coming together for better or for worse, hopefully enduring, and intimate to the degree of being sacred. The association promotes a way of life, not causes; a harmony in living, not political faiths; a bilateral loyalty, not commercial or social projects. Yet it is an association for as noble a purpose as any involved in our prior decisions.

Reversed.

Mr. Justice Goldberg, whom the Chief Justice [Warren] and Mr. Justice Brennan join, concurring [in the Court's opinion]: My conclusion that liberty embraces the right of marital privacy though that right is not mentioned explicitly in the Constitution is supported both by numerous decisions and by the language and history of the Ninth Amendment, which reveal that the Framers believed that there are additional fundamental rights, protected from governmental infringement, which exist alongside those fundamental rights specifically mentioned in the first eight amendments.

The Ninth Amendment reads, "The enumeration in the Constitution, of certain rights, shall not be construed to deny or disparage others retained by the people." It was proffered to quiet expressed fears that a bill of specifically enumerated rights could not be sufficiently broad to cover all essential rights and that the specific mention of certain rights would be interpreted as a denial that others were protected.

In determining which rights are fundamental, judges are not left at large to decide cases in light of their personal and private notions. Rather, they must look to the "traditions and collective conscience of our people" to determine whether a principle is "so rooted there as to be ranked as fundamental." The entire fabric of the Constitution and the purposes that clearly underlie its specific guarantees demonstrate that the rights to marital privacy and to marry and raise a family are of similar order and magnitude as the fundamental rights specifically protected. The logic of the dissents would sanction federal or state legislation that seems to me even more plainly unconstitutional than the statute before us. Surely the Government, absent a showing of a compelling subordinating state interest, could not decree that all husbands

and wives must be sterilized after two children have been born to them. If upon a showing of a slender basis of rationality, a law outlawing voluntary birth control by married persons is valid, then, by the same reasoning, a law requiring compulsory birth control also would seem to be valid. In my view, however, both types of law would unjustifiably intrude upon rights of marital privacy which are constitutionally protected.

In a long series of cases this Court has held that where fundamental personal liberties are involved, they may not be abridged by the States simply on a showing that a regulatory statute has some rational relationship to the effectuation of a proper state purpose. . . . The State, at most, argues that there is some rational relation between this statute and what is admittedly a legitimate subject of state concern—the discouraging of extramarital relations. It says that preventing the use of birth-control devices by married persons helps prevent the indulgence by some in such extramarital relations. The rationality of this justification is dubious, particularly in light of the admitted widespread availability to all persons in Connecticut, unmarried as well as married, of birth-control devices for the prevention of disease, as distinguished from the prevention of conception. But, in any event, it is clear that the state interest in safeguarding marital fidelity can be served by a more discriminately tailored statute, which does not, like the present one, sweep unnecessarily broadly, reaching far beyond the evil sought to be dealt with and intruding upon the privacy of all married couples. [Connecticut] does have statutes which prohibit adultery and fornication. These statutes demonstrate that means for achieving the same basic purpose of protecting marital fidelity are available to Connecticut without the need to "invade the area of protected freedoms." [Finally], it should be said of the Court's holding today that it in no way interferes with a State's proper regulation of sexual promiscuity or misconduct. In sum, I believe that the right of privacy in the marital relation is fundamental and basic—a personal right "retained by the people" within the meaning of the Ninth Amendment, [a right] which is protected by the Fourteenth Amendment from infringement by the [States].

Mr. Justice Black [joined by Justice Stewart], dissenting: The law is every bit as offensive to me as it is to my Brethren [who], reciting reasons why it is offensive to them, hold it unconstitutional. [But] I cannot [join] their conclusion. The Court talks about a constitutional "right of privacy". There are, of course, guarantees in certain specific constitutional provisions which are designed in part to protect privacy at certain times and places with respect to certain activities.

Such, for example, is the Fourth Amendment. But I think it belittles that Amendment to talk about it as though it protects nothing but "privacy." One of the most effective ways of diluting or expanding a constitutionally guaranteed right is to substitute for the crucial word or words of a constitutional guarantee another word or words, more or less flexible and more or less restricted in meaning. "Privacy" is a broad, abstract and ambiguous concept which can easily be shrunken in meaning but which can also easily be interpreted as a constitutional ban against many things other than searches and seizures. I get nowhere in this case by talk about a constitutional "right of privacy" as an emanation from one or more constitutional provisions. I like my privacy as well as the next one, but I am nevertheless compelled to admit that government has a right to invade it unless prohibited by some specific constitutional provision.

This brings me to the arguments made by my Brothers Harlan, White and Goldberg. I discuss the due process and Ninth Amendment arguments together because on analysis they turn out to be the same thing—merely using different words to claim the power to invalidate any legislative act which the judges find irrational, unreasonable or offensive. If formulas based on "natural justice" are to prevail, they require judges to determine what is or is not constitutional on the basis of their own appraisal of what laws are unwise or unnecessary. The power to make such decisions is of course that of a legislative body. [I] do not believe that we are granted power [to] measure constitutionality by our belief that the legislation is arbitrary, capricious or unreasonable, or accomplishes no justifiable purpose, or is offensive to our own notions of "civilized standards of conduct."

My Brother Goldberg has adopted the recent discovery that the Ninth Amendment as well as the

Due Process Clause can be used by this Court as authority to strike down all state legislation which this Court thinks violates "fundamental principles of liberty and justice," or is contrary to the "traditions and collective conscience of our people." He also states that in making decisions on this basis judges will not consider "their personal and private notions." One may ask how they can avoid considering them. [The Framers did not give this Court] veto powers over [lawmaking]. Nor does anything in the history of the Amendment offer any support for such a shocking doctrine. That Amendment was passed to assure the people that the [Constitution] was intended to limit the Federal Government to the powers granted expressly or by necessary implication. This fact is perhaps responsible for the peculiar phenomenon that for a period of a century and a half no serious suggestion was ever made that the Ninth Amendment, enacted to protect state powers against federal invasion, could be used as a weapon of federal power to prevent state legislatures from passing laws they consider appropriate to govern local affairs.

I realize that many good and able men have eloquently spoken and written [about] the duty of this Court to keep the Constitution in tune with the times. I reject that philosophy. The Constitution makers knew the need for change and provided for it [through the amendment process]. The Due Process Clause with an "arbitrary and capricious" or "shocking to the conscience" formula was liberally used by this Court to strike down economic legislation in the early decades of this century, threatening, many people thought, the tranquility and stability of the Nation. See, e.g., *Lochner.* That formula is no less dangerous when used to enforce this Court's views about personal rights than those about economic rights. I had thought that we had laid that formula, as a means for striking down state legislation, to rest. [Apparently] my Brethren have less quarrel with state economic regulations than former Justices of their persuasion had. But any limitation upon their using the natural law due process philosophy to strike down any state law, dealing with any activity whatever, will obviously be only self-imposed. The late Judge Learned Hand, after emphasizing his view that judges should not use the due process formula suggested in the concurring opinions today or any other formula like it to invalidate legislation offensive to their "personal preferences," made the statement, with which I fully agree, that: "For myself it would be most irksome to be ruled by a bevy of Platonic Guardians, even if I knew how to choose them, which I assuredly do not." So far as I am concerned, Connecticut's law as applied here is not forbidden by any provision of the Federal Constitution as that Constitution was written.

REVIEW AND DISCUSSION QUESTIONS

1. Explain Justice Douglas's reasoning.
2. How does Douglas answer those who say that the right to privacy is not mentioned in the Constitution?
3. Justice Goldberg agrees with the outcome of the case, but disagrees with the reasoning. Explain the differences between his opinion and Justice Douglas's.
4. How is the Ninth Amendment to be interpreted, according to Justice Goldberg?
5. Justice Black called the law "offensive," and yet he disagrees with the position that it is unconstitutional. On what basis does he dissent?
6. Would Mill have agreed that there is a moral right to privacy? If so, would that have given him reason to side with one of the decisions in this case? If so, which of the decisions most clearly mirrors Mill's philosophical position?

Homosexuality and the Right to Privacy

Lawrence v. Texas

Although the word "privacy" never appears in the U.S. Constitution, the Supreme Court had earlier held in *Griswold v. Connecticut* that the right exists in the "penumbras" of other rights, including rights against unreasonable searches and the right not to have soldiers quartered in homes. Although it relied on *Griswold* in later decisions, including *Roe v. Wade* (1973), in *Bowers v. Hardwick* (1986) the Court refused to extend the right to privacy to consensual homosexual conduct, upholding state laws that banned it. In the 2003 case of *Lawrence v. Texas,* however, the Court revisited its decision in *Bowers,* concluding that its earlier decision was an error.

Justice Kennedy delivered the opinion of the Court: Liberty protects the person from unwarranted government intrusions into a dwelling or other private places. In our tradition the State is not omnipresent in the home. And there are other spheres of our lives and existence, outside the home, where the State should not be a dominant presence. Freedom extends beyond spatial bounds. Liberty presumes an autonomy of self that includes freedom of thought, belief, expression, and certain intimate conduct. The instant case involves liberty of the person both in its spatial and more transcendent dimensions.

I

The question before the Court is the validity of a Texas statute making it a crime for two persons of the same sex to engage in . . . "deviate sexual intercourse, namely anal sex, with a member of the same sex (man)." . . .

II

We conclude the case should be resolved by determining whether the petitioners were free as adults to engage in the private conduct in the exercise of their liberty under the Due Process Clause of the Fourteenth Amendment to the Constitution. For this inquiry we deem it necessary to reconsider the Court's holding in *Bowers*. . . .

In *Griswold* the Court invalidated a state law prohibiting the use of drugs or devices of contraception and counseling or aiding and abetting the use of contraceptives. The Court described the protected interest as a right to privacy and placed emphasis on the marriage relation and the protected space of the marital bedroom.

After *Griswold* it was established that the right to make certain decisions regarding sexual conduct extends beyond the marital relationship. In *Eisenstadt v. Baird* the Court invalidated a law prohibiting the distribution of contraceptives to unmarried persons. The case was decided under the Equal Protection Clause, but with respect to unmarried persons, the Court went on to state the fundamental proposition that the law impaired the exercise of their personal rights. It quoted from the statement of the Court of Appeals finding the law to be in conflict with fundamental human rights, and it followed with this statement of its own:

> It is true that in *Griswold* the right of privacy in question inhered in the marital relationship. . . . If the right of privacy means anything, it is the right of the *individual,* married or single, to be free from unwarranted governmental intrusion into matters so fundamentally affecting a person as the decision whether to bear or beget a child.

The opinions in *Griswold* and *Eisenstadt* were part of the background for the decision in *Roe v. Wade*. . . . *Roe* recognized the right of a woman to make certain fundamental decisions affecting her destiny and confirmed once more that the protection

---

59 U.S. 558 (2003). Most citations omitted.

of liberty under the Due Process Clause has a substantive dimension of fundamental significance in defining the rights of the person. . . .

This was the state of the law with respect to some of the most relevant cases when the Court considered *Bowers v. Hardwick*. . . .

The Court began its substantive discussion in *Bowers* as follows: "The issue presented is whether the Federal Constitution confers a fundamental right upon homosexuals to engage in sodomy and hence invalidates the laws of the many States that still make such conduct illegal and have done so for a very long time." That statement, we now conclude, discloses the Court's own failure to appreciate the extent of the liberty at stake. To say that the issue in *Bowers* was simply the right to engage in certain sexual conduct demeans the claim the individual pút forward, just as it would demean a married couple were it to be said marriage is simply about the right to have sexual intercourse. The laws involved in *Bowers* and there are, to be sure, statutes that purport to do no more than prohibit a particular sexual act. Their penalties and purposes, though, have more far-reaching consequences, touching upon the most private human conduct, sexual behavior, and in the most private of places, the home. The statutes do seek to control a personal relationship that, whether or not entitled to formal recognition in the law, is within the liberty of persons to choose without being punished as criminals.

This, as a general rule, should counsel against attempts by the State, or a court, to define the meaning of the relationship or to set its boundaries absent injury to a person or abuse of an institution the law protects. It suffices for us to acknowledge that adults may choose to enter upon this relationship in the confines of their homes and their own private lives and still retain their dignity as free persons. When sexuality finds overt expression in intimate conduct with another person, the conduct can be but one element in a personal bond that is more enduring. The liberty protected by the Constitution allows homosexual persons the right to make this choice.

Having misapprehended the claim of liberty there presented to it, and thus stating the claim to be whether there is a fundamental right to engage in consensual sodomy, the *Bowers* Court said: "Proscriptions against that conduct have ancient roots." In academic writings, and in many of the scholarly amicus briefs filed to assist the Court in this case, there are fundamental criticisms of [this statement.] We need not enter this debate in the attempt to reach a definitive historical judgment, but the following considerations counsel against adopting the definitive conclusions upon which *Bowers* placed such reliance.

At the outset it should be noted that there is no longstanding history in this country of laws directed at homosexual conduct as a distinct matter. . . . Thus early American sodomy laws were not directed at homosexuals as such but instead sought to prohibit nonprocreative sexual activity more generally. This does not suggest approval of homosexual conduct. It does tend to show that this particular form of conduct was not thought of as a separate category from like conduct between heterosexual persons.

Laws prohibiting sodomy do not seem to have been enforced against consenting adults acting in private. A substantial number of sodomy prosecutions and convictions for which there are surviving records were for predatory acts against those who could not or did not consent, as in the case of a minor or the victim of an assault. As to these, one purpose for the prohibitions was to ensure there would be no lack of coverage if a predator committed a sexual assault that did not constitute rape as defined by the criminal law. Thus the model sodomy indictments presented in a 19th-century treatise addressed the predatory acts of an adult man against a minor girl or minor boy. Instead of targeting relations between consenting adults in private, 19th-century sodomy prosecutions typically involved relations between men and minor girls or minor boys, relations between adults involving force, relations between adults implicating disparity in status, or relations between men and animals. . . .

It was not until the 1970's that any State singled out same-sex relations for criminal prosecution, and only nine States have done so. Post-*Bowers* even some of these States did not adhere to the policy of suppressing homosexual conduct. Over the course of the last decades. States with same-sex prohibitions have moved toward abolishing them. . . .

It must be acknowledged, of course, that the Court in *Bowers* was making the broader point that for centuries there have been powerful voices to condemn homosexual conduct as immoral. The condemnation has been shaped by religious beliefs, conceptions of right and acceptable behavior, and respect for the traditional family. For many persons these are not trivial concerns but profound and deep convictions accepted as ethical and moral principles to which they aspire and which thus determine the course of their lives. These considerations do not answer the question before us, however. The issue is whether the majority may use the power of the State to enforce these views on the whole society through operation of the criminal law. . . .

. . . In *Planned Parenthood of Southeastern Pa. v. Casey* the Court reaffirmed the substantive force of the liberty protected by the Due Process Clause. The *Casey* decision again confirmed that our laws and tradition afford constitutional protection to personal decisions relating to marriage, procreation, contraception, family relationships, child rearing, and education. In explaining the respect the Constitution demands for the autonomy of the person in making these choices, we stated as follows:

> These matters, involving the most intimate and personal choices a person may make in a lifetime, choices central to personal dignity and autonomy, are central to the liberty protected by the Fourteenth Amendment. At the heart of liberty is the right to define one's own concept of existence, of meaning, of the universe, and of the mystery of human life. Beliefs about these matters could not define the attributes of personhood were they formed under compulsion of the State.

Persons in a homosexual relationship may seek autonomy for these purposes, just as heterosexual persons do. The decision in *Bowers* would deny them this right. . . .

Equality of treatment and the due process right to demand respect for conduct protected by the substantive guarantee of liberty are linked in important respects, and a decision on the latter point advances both interests. If protected conduct is made criminal and the law which does so remains unexamined for its substantive validity, its stigma might remain even if it were not enforceable as drawn for equal protection reasons. When homosexual conduct is made criminal by the law of the State, that declaration in and of itself is an invitation to subject homosexual persons to discrimination both in the public and in the private spheres. The central holding of *Bowers* has been brought in question by this case, and it should be addressed. Its continuance as precedent demeans the lives of homosexual persons.

The stigma this criminal statute imposes, moreover, is not trivial. The offense, to be sure, is but a class C misdemeanor, a minor offense in the Texas legal system. Still, it remains a criminal offense with all that imports for the dignity of the persons charged. The petitioners will bear on their record the history of their criminal convictions. . . .

To the extent *Bowers* relied on values we share with a wider civilization, it should be noted that the reasoning and holding in *Bowers* have been rejected elsewhere. The European Court of Human Rights has followed not *Bowers* but its own decision in *Dudgeon v. United Kingdom.* Other nations, too, have taken action consistent with an affirmation of the protected right of homosexual adults to engage in intimate, consensual conduct. The right the petitioners seek in this case has been accepted as an integral part of human freedom in many other countries. There has been no showing that in this country the governmental interest in circumscribing personal choice is somehow more legitimate or urgent. . . .

. . . Indeed, there has been no individual or societal reliance on *Bowers* of the sort that could counsel against overturning its holding once there are compelling reasons to do so. *Bowers* itself causes uncertainty, for the precedents before and after its issuance contradict its central holding.

The rationale of *Bowers* does not withstand careful analysis. In his dissenting opinion in *Bowers* Justice Stevens came to these conclusions:

> Our prior cases make two propositions abundantly clear. First, the fact that the governing majority in a State has traditionally viewed a particular practice as immoral is not a sufficient reason for upholding a law prohibiting the practice; neither history nor tradition could save a law prohibiting miscegenation from constitutional attack. Second, individual decisions by married persons, concerning the intimacies of their physical

> relationship, even when not intended to produce offspring, are a form of "liberty" protected by the Due Process Clause of the fourteenth Amendment. Moreover, this protection extends to intimate choices by unmarried as well as married persons.

Justice Stevens' analysis, in our view, should have been controlling in *Bowers* and should control here.

Bowers was not correct when it was decided, and it is not correct today. It ought not to remain binding precedent. *Bowers v. Hardwick* should be and now is overruled.

The present case does not involve minors. It does not involve persons who might be injured or coerced or who are situated in relationships where consent might not easily be refused. It does not involve public conduct or prostitution. It does not involve whether the government must give formal recognition to any relationship that homosexual persons seek to enter. The case does involve two adults who, with full and mutual consent from each other, engaged in sexual practices common to a homosexual lifestyle. The petitioners are entitled to respect for their private lives. The State cannot demean their existence or control their destiny by making their private sexual conduct a crime. Their right to liberty under the Due Process Clause gives them the full right to engage in their conduct without intervention of the government. The Texas statute furthers no legitimate state interest which can justify its intrusion into the personal and private life of the individual.

Had those who drew and ratified the Due Process Clauses of the Fifth Amendment or the Fourteenth Amendment known the components of liberty in its manifold possibilities, they might have been more specific. They did not presume to have this insight. They knew times can blind us to certain truths and later generations can see that laws once thought necessary and proper in fact serve only to oppress. As the Constitution endures, persons in every generation can invoke its principles in their own search for greater freedom. . . .

JUSTICE O'CONNOR, concurring in the judgment: . . . This case raises a different issue than *Bowers:* whether, under the Equal Protection Clause, moral disapproval is a legitimate state interest to justify by itself a statute that bans homosexual sodomy, but not heterosexual sodomy. It is not. Moral disapproval of this group, like a bare desire to harm the group, is an interest that is insufficient to satisfy rational basis review under the Equal Protection Clause. Indeed, we have never held that moral disapproval, without any other asserted state interest, is a sufficient rationale under the Equal Protection Clause to justify a law that discriminates among groups of persons.

Moral disapproval of a group cannot be a legitimate governmental interest under the Equal Protection Clause because legal classifications must not be "drawn for the purpose of disadvantaging the group burdened by the law." Texas' invocation of moral disapproval as a legitimate state interest proves nothing more than Texas' desire to criminalize homosexual sodomy. But the Equal Protection Clause prevents a State from creating "a classification of persons" undertaken for its own sake. And because Texas so rarely enforces its sodomy law as applied to private, consensual acts, the law serves more as a statement of dislike and disapproval against homosexuals than as a tool to stop criminal behavior. . . .

A State can of course assign certain consequences to a violation of its criminal law. But the State cannot single out one identifiable class of citizens for punishment that does not apply to everyone else, with moral disapproval as the only asserted state interest for the law. The Texas sodomy statute subjects homosexuals to "a life-long penalty and stigma". . . .

JUSTICE SCALIA, with whom THE CHIEF JUSTICE and JUSTICE THOMAS join, dissenting: [I] begin with the Court's surprising readiness to reconsider a decision rendered a mere 17 years ago in *Bowers v. Hardwick.* I do not myself believe in rigid adherence to stare decisis in constitutional cases; but I do believe that we should be consistent rather than manipulative in invoking the doctrine. Today's opinions in support of reversal do not bother to distinguish—or indeed, even bother to mention—the paean to stare decisis coauthored by three Members of today's majority in *Planned Parenthood v. Casey.* . . .

Today's approach to stare decisis invites us to overrule an erroneously decided precedent

(including an "intensely divisive" decision) if: (1) its foundations have been "eroded" by subsequent decisions; (2) it has been subject to "substantial and continuing" criticism; and (3) it has not induced "individual or societal reliance" that counsels against overturning. The problem is that *Roe* itself—which today's majority surely has no disposition to overrule—satisfies these conditions to at least the same degree as *Bowers*. . . .

Bowers, the Court says, has been subject to "substantial and continuing [criticism], disapproving of its reasoning in all respects, not just as to its historical assumptions." [Of] course, *Roe* too (and by extension *Casey*) had been (and still is) subject to unrelenting criticism, including criticism from the two commentators cited by the Court today.

That leaves, to distinguish the rock-solid, unamendable disposition of *Roe* from the readily overrulable *Bowers,* only the third factor. "[T]here has been," the Court says, "no individual or societal reliance on *Bowers* of the sort that could counsel against overturning its holding. . . ." It seems to me that the "societal reliance" on the principles confirmed in *Bowers* and discarded today has been overwhelming. Countless judicial decisions and legislative enactments have relied on the ancient proposition that a governing majority's belief that certain sexual behavior is "immoral and unacceptable" constitutes a rational basis for regulation. . . . State laws against bigamy, same-sex marriage, adult incest, prostitution, masturbation, adultery, fornication, bestiality, and obscenity are likewise sustainable only in light of *Bowers*' validation of laws based on moral choices. Every single one of these laws is called into question by today's decision; the Court makes no effort to cabin the scope of its decision to exclude them from its holding. The impossibility of distinguishing homosexuality from other traditional "morals" offenses is precisely why *Bowers* rejected the rational-basis challenge. . . .

What a massive disruption of the current social order, therefore, the overruling of *Bowers* entails. Not so the overruling of *Roe,* which would simply have restored the regime that existed for centuries before 1973, in which the permissibility of and restrictions upon abortion were determined legislatively State by State. . . .

To tell the truth, it does not surprise me, and should surprise no one, that the Court has chosen today to revise the standards of stare decisis set forth in *Casey.* It has thereby exposed *Casey*'s extraordinary deference to precedent for the result-oriented expedient that it is. . . .

Texas Penal Code Ann. § 21.06(a) (2003) undoubtedly imposes constraints on liberty. So do laws prohibiting prostitution, recreational use of heroin, and, for that matter, working more than 60 hours per week in a bakery. But there is no right to "liberty" under the Due Process Clause, though today's opinion repeatedly makes that claim. The Fourteenth Amendment expressly allows States to deprive their citizens of "liberty," so long as "due process of law" is provided:

> No state shall . . . "deprive any person of life, liberty, or property, *without due process of law.*" Amdt. 14 (emphasis added).

Our opinions applying the doctrine known as "substantive due process" hold that the Due Process Clause prohibits States from infringing fundamental liberty interests, unless the infringement is narrowly tailored to serve a compelling state interest. We have held repeatedly, in cases the Court today does not overrule, that only fundamental rights qualify for this so-called "heightened scrutiny" protection—that is, rights which are "deeply rooted in this Nation's history and tradition." All other liberty interests may be abridged or abrogated pursuant to a validly enacted state law if that law is rationally related to a legitimate state interest.

The Court's description of "the state of the law" at the time of *Bowers* only confirms that *Bowers* was right. The Court points to Griswold v. Connecticut. But that case expressly disclaimed any reliance on the doctrine of "substantive due process," and grounded the so-called "right to privacy" in penumbras of constitutional provisions other than the Due Process Clause. . . .

After discussing the history of antisodomy laws, the Court proclaims that, "it should be noted that there is no longstanding history in this country of laws directed at homosexual conduct as a distinct matter." This observation in no way casts into doubt the "definitive [historical] conclusion," on which *Bowers* relied: that our Nation has a longstanding history of laws prohibiting sodomy in

general—regardless of whether it was performed by same-sex or opposite-sex couples. . . .

. . . Whether homosexual sodomy was prohibited by a law targeted at same-sex sexual relations or by a more general law prohibiting both homosexual and heterosexual sodomy, the only relevant point is that it was criminalized—which suffices to establish that homosexual sodomy is not a right "deeply rooted in our Nation's history and tradition." The Court today agrees that homosexual sodomy was criminalized and thus does not dispute the facts on which *Bowers* actually relied.

Next the Court makes the claim, again unsupported by any citations, that "[l]aws prohibiting sodomy do not seem to have been enforced against consenting adults acting in private." The key qualifier here is "acting in private"—since the Court admits that sodomy laws were enforced against consenting adults (although the Court contends that prosecutions were "infrequent"). I do not know what "acting in private" means; surely consensual sodomy, like heterosexual intercourse, is rarely performed on stage. . . . *Bowers*' conclusion that homosexual sodomy is not a fundamental right "deeply rooted in this Nation's history and tradition" is utterly unassailable.

Realizing that fact, the Court instead says: "[W]e think that our laws and traditions in the past half century are of most relevance here. These references show *an emerging awareness* that liberty gives substantial protection to adult persons in deciding how to conduct their private lives *in matters pertaining to sex*" (emphasis added). Apart from the fact that such an "emerging awareness" does not establish a "fundamental right," the statement is factually false. States continue to prosecute all sorts of crimes by adults "in matters pertaining to sex": prostitution, adult incest, adultery, obscenity, and child pornography. Sodomy laws, too, have been enforced "in the past half century," in which there have been 134 reported cases involving prosecutions for consensual, adult, homosexual sodomy. . . .

In any event, an "emerging awareness" is by definition not "deeply rooted in this Nation's history and tradition[s]," as we have said "fundamental right" status requires. Constitutional entitlements do not spring into existence because some States choose to lessen or eliminate criminal sanctions on certain behavior. Much less do they spring into existence, as the Court seems to believe, because foreign nations decriminalize conduct. . . . The Court's discussion of these foreign views (ignoring, of course, the many countries that have retained criminal prohibitions on sodomy) is therefore meaningless dicta. . . .

I turn now to the ground on which the Court squarely rests its holding: the contention that there is no rational basis for the law here under attack. This proposition is so out of accord with our jurisprudence—indeed, with the jurisprudence of any society we know—that it requires little discussion.

The Texas statute undeniably seeks to further the belief of its citizens that certain forms of sexual behavior are "immoral and unacceptable"—the same interest furthered by criminal laws against fornication, bigamy, adultery, adult incest, bestiality, and obscenity. *Bowers* held that this was a legitimate state interest. The Court today reaches the opposite conclusion. The Texas statute, it says, "furthers no legitimate state interest which can justify its intrusion into the personal and private life of the individual." The Court embraces instead Justice Stevens' declaration in his *Bowers* dissent, that "the fact that the governing majority in a State has traditionally viewed a particular practice as immoral is not a sufficient reason for upholding a law prohibiting the practice." This effectively decrees the end of all morals legislation. If, as the Court asserts, the promotion of majoritarian sexual morality is not even a legitimate state interest, none of the above-mentioned laws can survive rational-basis review. . . .

Today's opinion is the product of a Court, which is the product of a law-profession culture, that has largely signed on to the so-called homosexual agenda, by which I mean the agenda promoted by some homosexual activists directed at eliminating the moral opprobrium that has traditionally attached to homosexual conduct. . . .

One of the most revealing statements in today's opinion is the Court's grim warning that the criminalization of homosexual conduct is "an invitation to subject homosexual persons to discrimination both in the public and in the private spheres." It is clear from this that the Court has taken sides in the culture war, departing from its role of assuring, as

neutral observer, that the democratic rules of engagement are observed. Many Americans do not want persons who openly engage in homosexual conduct as partners in their business, as scoutmasters for their children, as teachers in their children's schools, or as boarders in their home. They view this as protecting themselves and their families from a lifestyle that they believe to be immoral and destructive. The Court views it as "discrimination" which it is the function of our judgments to deter. So imbued is the Court with the law profession's anti-anti-homosexual culture, that it is seemingly unaware that the attitudes of that culture are not obviously "mainstream"; that in most States what the Court calls "discrimination" against those who engage in homosexual acts is perfectly legal; that proposals to ban such "discrimination" under Title VII have repeatedly been rejected by Congress; that in some cases such "discrimination" is mandated by federal statute, see 10 U.S.C. § 654(b)(1) (mandating discharge from the armed forces of any service member who engages in or intends to engage in homosexual acts); and that in some cases such "discrimination" is a constitutional right, see *Boy Scouts of America v. Dale.*

Let me be clear that I have nothing against homosexuals, or any other group, promoting their agenda through normal democratic means. Social perceptions of sexual and other morality change over time, and every group has the right to persuade its fellow citizens that its view of such matters is the best. That homosexuals have achieved some success in that enterprise is attested to by the fact that Texas is one of the few remaining States that criminalize private, consensual homosexual acts. But persuading one's fellow citizens is one thing, and imposing one's views in absence of democratic majority will is something else. I would no more require a State to criminalize homosexual acts—or, for that matter, display any moral disapprobation of them—than I would forbid it to do so. What Texas has chosen to do is well within the range of traditional democratic action, and its hand should not be stayed through the invention of a brand-new "constitutional right" by a Court that is impatient of democratic change. . . .

One of the benefits of leaving regulation of this matter to the people rather than to the courts is that the people, unlike judges, need not carry things to their logical conclusion. The people may feel that their disapprobation of homosexual conduct is strong enough to disallow homosexual marriage, but not strong enough to criminalize private homosexual acts—and may legislate accordingly. The Court today pretends that it possesses a similar freedom of action, so that we need not fear judicial imposition of homosexual marriage. . . . Do not believe it. . . . Today's opinion dismantles the structure of constitutional law that has permitted a distinction to be made between heterosexual and homosexual unions, insofar as formal recognition in marriage is concerned. If moral disapprobation of homosexual conduct is "no legitimate state interest" for purposes of proscribing that conduct and if, as the Court coos (casting aside all pretense of neutrality), "[w]hen sexuality finds overt expression in intimate conduct with another person, the conduct can be but one element in a personal bond that is more enduring"; what justification could there possibly be for denying the benefits of marriage to homosexual couples exercising "[t]he liberty protected by the Constitution"? Surely not the encouragement of procreation, since the sterile and the elderly are allowed to marry. This case "does not involve" the issue of homosexual marriage only if one entertains the belief that principle and logic have nothing to do with the decisions of this Court. . . .

Justice Thomas, dissenting: I join Justice Scalia's dissenting opinion. I write separately to note that the law before the Court today "is . . . uncommonly silly." . . . If I were a member of the Texas Legislature, I would vote to repeal it. Punishing someone for expressing his sexual preference through noncommercial consensual conduct with another adult does not appear to be a worthy way to expend valuable law enforcement resources.

Notwithstanding this, I recognize that as a member of this Court I am not empowered to help petitioners and others similarly situated. My duty, rather, is to "decide cases 'agreeably to the Constitution and laws of the United States.'" And, just like Justice Stewart, I "can find [neither in the Bill of Rights nor any other part of the Constitution a] general right of privacy," or as the Court terms it today, the "liberty of the person both in its spatial and more transcendent dimensions."

REVIEW AND DISCUSSION QUESTIONS

1. What was the Court's reasoning in *Bowers*?
2. Why did the Court reject that reasoning in this case, holding instead that the Texas statute violated individual rights? What right(s) were at stake, according to Justice Kennedy's majority opinion?
3. Justice Scalia argues that the Court's logic in this case would also lead it to reject *Roe v. Wade*. Why does he believe that?
4. Summarize Justice Scalia's argument that this statute does not involve a violation of the right to "substantive due process."
5. It is sometimes said that judges decide cases like this based on their own beliefs about the morality of the conduct in question. Explain whether you think one or both sides are guilty of that charge. Do you think that judges *should* try to avoid letting such beliefs influence their decisions?
6. Discuss this case in light of the debate between Scalia and Dworkin in Section 6 about how the Constitution should be interpreted.

Same-Sex Marriage

Goodridge v. Massachusetts Department of Public Health

This case, decided in 2003, raises the question whether laws that limit marriage to a man and a woman are constitutional. The majority of the Massachusetts Court found that the law violates people's fundamental rights, which means in effect that there must be a strong, compelling need or interest on the part of the state for the law. The dissenters denied there is a fundamental right at stake and argued that there need only be a "rational basis" for the law. Besides the substantive issue of the nature and purposes of marriage, the Court also addresses questions about judicial power and the extent to which courts should defer to legislative bodies when controversial moral and social issues are being addressed. (The "Department" here refers to the State of Massachusetts, which was defending the constitutionality of the law.)

MARSHALL, C.J: Marriage is a vital social institution. The exclusive commitment of two individuals to each other nurtures love and mutual support; it brings stability to our society. For those who choose to marry, and for their children, marriage provides an abundance of legal, financial, and social benefits. In return it imposes weighty legal, financial, and social obligations. The question before us is whether, consistent with the Massachusetts Constitution, the Commonwealth may deny the protections, benefits, and obligations conferred by civil marriage to two individuals of the same sex who wish to marry. We conclude that it may not. The Massachusetts Constitution affirms the dignity and equality of all individuals. It forbids the creation of second-class citizens. In reaching our conclusion we have given full deference to the arguments made by the Commonwealth. But it has failed to identify any constitutionally adequate reason for denying civil marriage to same-sex couples.

We are mindful that our decision marks a change in the history of our marriage law. Many people hold deep-seated religious, moral, and ethical convictions that marriage should be limited to the union of one man and one woman, and that homosexual conduct is immoral. Many hold equally strong religious, moral, and ethical convictions that same-sex couples are entitled to be married,

440 Mass. 309 (2003). Most citations omitted.

and that homosexual persons should be treated no differently than their heterosexual neighbors. Neither view answers the question before us. Our concern is with the Massachusetts Constitution as a charter of governance for every person properly within its reach. "Our obligation is to define the liberty of all, not to mandate our own moral code." *Lawrence v. Texas* (2003), quoting *Planned Parenthood of Southeastern Pa. v. Casey* (1992).

Barred access to the protections, benefits, and obligations of civil marriage, a person who enters into an intimate, exclusive union with another of the same sex is arbitrarily deprived of membership in one of our community's most rewarding and cherished institutions. That exclusion is incompatible with the constitutional principles of respect for individual autonomy and equality under law.

Although the plaintiffs refer in passing to "the marriage statutes," they focus, quite properly, on G. L. c. 207, the marriage licensing statute, which controls entry into civil marriage. As a preliminary matter, we summarize the provisions of that law.

General Laws c. 207 is both a gatekeeping and a public records statute. The gatekeeping provisions forbid marriage of individuals within certain degrees of consanguinity, and polygamous marriages. They prohibit marriage if one of the parties has communicable syphilis and restrict the circumstances in which a person under eighteen years of age may marry. The plaintiffs argue that because nothing in that licensing law specifically prohibits marriages between persons of the same sex, we may interpret the statute to permit "qualified same sex couples" to obtain marriage licenses, thereby avoiding the question whether the law is constitutional. This claim lacks merit.

We interpret statutes to carry out the Legislature's intent, determined by the words of a statute interpreted according to "the ordinary and approved usage of the language." The everyday meaning of "marriage" is "[t]he legal union of a man and woman as husband and wife," *Black's Law Dictionary* 986 (7th ed. 1999), and the plaintiffs do not argue that the term "marriage" has ever had a different meaning under Massachusetts law.

The intended scope of G. L. c. 207 is also evident in its consanguinity provisions. Sections 1 and 2 of G. L. c. 207 prohibit marriages between a man and certain female relatives and a woman and certain male relatives, but are silent as to the consanguinity of male-male or female-female marriage applicants. The only reasonable explanation is that the Legislature did not intend that same-sex couples be licensed to marry. We conclude, as did the judge, that G. L. c. 207 may not be construed to permit same-sex couples to marry.

The larger question is whether, as the department claims, government action that bars same-sex couples from civil marriage constitutes a legitimate exercise of the State's authority to regulate conduct, or whether, as the plaintiffs claim, this categorical marriage exclusion violates the Massachusetts Constitution. We have recognized the long-standing statutory understanding, derived from the common law, that "marriage" means the lawful union of a woman and a man. But that history cannot and does not foreclose the constitutional question.

The plaintiffs' claim that the marriage restriction violates the Massachusetts Constitution can be analyzed in two ways. Does it offend the Constitution's guarantees of equality before the law? Or do the liberty and due process provisions of the Massachusetts Constitution secure the plaintiffs' right to marry their chosen partner? In matters implicating marriage, family life, and the upbringing of children, the two constitutional concepts frequently overlap, as they do here.

We begin by considering the nature of civil marriage itself. Simply put, the government creates civil marriage. In Massachusetts, civil marriage is, and since pre-Colonial days has been, precisely what its name implies: a wholly secular institution.

In a real sense, there are three partners to every civil marriage: two willing spouses and an approving State. While only the parties can mutually assent to marriage, the terms of the marriage—who may marry and what obligations, benefits, and liabilities attach to civil marriage—are set by the Commonwealth.

Without question, civil marriage enhances the "welfare of the community." It is a "social institution of the highest importance." Civil marriage anchors an ordered society by encouraging stable relationships over transient ones. It is central to the way the Commonwealth identifies individuals, provides for the orderly distribution of property, ensures that children and adults are cared for and supported whenever possible from private rather

than public funds, and tracks important epidemiological and demographic data.

Marriage also bestows enormous private and social advantages on those who choose to marry. Civil marriage is at once a deeply personal commitment to another human being and a highly public celebration of the ideals of mutuality, companionship, intimacy, fidelity, and family. "It is an association that promotes a way of life, not causes; a harmony in living, not political faiths; a bilateral loyalty, not commercial or social projects." *Griswold v. Connecticut* (1965). Because it fulfils yearnings for security, safe haven, and connection that express our common humanity, civil marriage is an esteemed institution, and the decision whether and whom to marry is among life's momentous acts of self-definition.

Tangible as well as intangible benefits flow from marriage. With no attempt to be comprehensive, we note that some of the statutory benefits conferred by the Legislature on those who enter into civil marriage include, as to property: joint Massachusetts income tax filing; tenancy by the entirety (a form of ownership that provides certain protections against creditors and allows for the automatic descent of property to the surviving spouse without probate); extension of the benefit of the homestead protection (securing up to $300,000 in equity from creditors) to one's spouse and children; automatic rights to inherit the property of a deceased spouse who does not leave a will; the rights of elective share and of dower (which allow surviving spouses certain property rights where the decedent spouse has not made adequate provision for the survivor in a will); entitlement to wages owed to a deceased employee; eligibility to continue certain businesses of a deceased spouse; the right to share the medical policy of one's spouse; preferential options under the Commonwealth's pension system; preferential benefits in the Commonwealth's medical program, financial protections for spouses of certain Commonwealth employees.

It is undoubtedly for these concrete reasons, as well as for its intimately personal significance, that civil marriage has long been termed a "civil right." The United States Supreme Court has described the right to marry as "of fundamental importance for all individuals" and as "part of the fundamental 'right of privacy' implicit in the Fourteenth Amendment's Due Process Clause." Without the right to marry—or more properly, the right to choose to marry—one is excluded from the full range of human experience and denied full protection of the laws.

The department posits three legislative rationales for prohibiting same-sex couples from marrying: (1) providing a "favorable setting for procreation"; (2) ensuring the optimal setting for child rearing, which the department defines as "a two-parent family with one parent of each sex"; and (3) preserving scarce State and private financial resources. We consider each in turn.

Our laws of civil marriage do not privilege procreative heterosexual intercourse between married people above every other form of adult intimacy and every other means of creating a family. The law contains no requirement that the applicants for a marriage license attest to their ability or intention to conceive children by coitus. Fertility is not a condition of marriage, nor is it grounds for divorce. People who have never consummated their marriage, and never plan to, may be and stay married.

Moreover, the Commonwealth affirmatively facilitates bringing children into a family regardless of whether the intended parent is married or unmarried, whether the child is adopted or born into a family, whether assistive technology was used to conceive the child, and whether the parent or her partner is heterosexual, homosexual, or bisexual. If procreation were a necessary component of civil marriage, our statutes would draw a tighter circle around the permissible bounds of nonmarital child bearing and the creation of families by noncoital means. The attempt to isolate procreation as "the source of a fundamental right to marry," post at (Cordy, J., dissenting), overlooks the integrated way in which courts have examined the complex and overlapping realms of personal autonomy, marriage, family life, and child rearing. Our jurisprudence recognizes that, in these nuanced and fundamentally private areas of life, such a narrow focus is inappropriate.

The department's first stated rationale, equating marriage with unassisted heterosexual procreation, shades imperceptibly into its second: that confining marriage to opposite-sex couples ensures that children are raised in the "optimal" setting. Protecting the welfare of children is a paramount State policy.

The department has offered no evidence that forbidding marriage to people of the same sex will increase the number of couples choosing to enter into opposite-sex marriages in order to have and raise children. There is thus no rational relationship between the marriage statute and the Commonwealth's proffered goal of protecting the "optimal" child rearing unit. Moreover, the department readily concedes that people in same-sex couples may be "excellent" parents. These couples (including four of the plaintiff couples) have children for the reasons others do—to love them, to care for them, to nurture them. But the task of child rearing for same-sex couples is made infinitely harder by their status as outliers to the marriage laws. Same-sex couples must undergo the sometimes lengthy and intrusive process of second-parent adoption to establish their joint parentage. While the enhanced income provided by marital benefits is an important source of security and stability for married couples and their children, those benefits are denied to families headed by same-sex couples.

It cannot be rational under our laws, and indeed it is not permitted, to penalize children by depriving them of State benefits because the State disapproves of their parents' sexual orientation.

The third rationale advanced by the department is that limiting marriage to opposite-sex couples furthers the Legislature's interest in conserving scarce State and private financial resources. [But] an absolute statutory ban on same-sex marriage bears no rational relationship to the goal of economy.

Here, the plaintiffs seek only to be married, not to undermine the institution of civil marriage. They do not want marriage abolished. They do not attack the binary nature of marriage, the consanguinity provisions, or any of the other gate-keeping provisions of the marriage licensing law. Recognizing the right of an individual to marry a person of the same sex will not diminish the validity or dignity of opposite-sex marriage, any more than recognizing the right of an individual to marry a person of a different race devalues the marriage of a person who marries someone of her own race. If anything, extending civil marriage to same-sex couples reinforces the importance of marriage to individuals and communities. That same-sex couples are willing to embrace marriage's solemn obligations of exclusivity, mutual support, and commitment to one another is a testament to the enduring place of marriage in our laws and in the human spirit.

It has been argued that, due to the State's strong interest in the institution of marriage as a stabilizing social structure, only the Legislature can control and define its boundaries. Accordingly, our elected representatives legitimately may choose to exclude same-sex couples from civil marriage in order to assure all citizens of the Commonwealth that (1) the benefits of our marriage laws are available explicitly to create and support a family setting that is, in the Legislature's view, optimal for child rearing, and (2) the State does not endorse gay and lesbian parenthood as the equivalent of being raised by one's married biological parents. These arguments miss the point. The Massachusetts Constitution requires that legislation meet certain criteria and not extend beyond certain limits. It is the function of courts to determine whether these criteria are met and whether these limits are exceeded. In most instances, these limits are defined by whether a rational basis exists to conclude that legislation will bring about a rational result. The Legislature in the first instance, and the courts in the last instance, must ascertain whether such a rational basis exists.

. . . Alarms about the imminent erosion of the "natural" order of marriage were sounded over the demise of anti-miscegenation laws, the expansion of the rights of married women, and the introduction of "no-fault" divorce. Marriage has survived all of these transformations, and we have no doubt that marriage will continue to be a vibrant and revered institution.

The marriage ban works a deep and scarring hardship on a very real segment of the community for no rational reason. The absence of any reasonable relationship between, on the one hand, an absolute disqualification of same-sex couples who wish to enter into civil marriage and, on the other, protection of public health, safety, or general welfare, suggests that the marriage restriction is rooted in persistent prejudices against persons who are (or who are believed to be) homosexual.

CORDY, J. (dissenting, with whom SPINA and SOSMAN, J.J., join): I respectfully dissent. Although it may be desirable for many reasons to extend to same-sex couples the benefits and

burdens of civil marriage (and the plaintiffs have made a powerfully reasoned case for that extension), that decision must be made by the Legislature, not the court.

The Massachusetts marriage statute does not impair the exercise of a recognized fundamental right, or discriminate on the basis of sex in violation of the equal rights amendment to the Massachusetts Constitution. Consequently, it is subject to review only to determine whether it satisfies the rational basis test. Because a conceivable rational basis exists upon which the Legislature could conclude that the marriage statute furthers the legitimate State purpose of ensuring, promoting, and supporting an optimal social structure for the bearing and raising of children, it is a valid exercise of the State's police power.

The plaintiffs ground their contention that they have a fundamental right to marry a person of the same sex in a long line of Supreme Court decisions that discuss the importance of marriage. Supreme Court cases that have described marriage or the right to marry as "fundamental" have focused primarily on the underlying interest of every individual in procreation, which, historically, could only legally occur within the construct of marriage because sexual intercourse outside of marriage was a criminal act. Supreme Court cases recogniz[e] a right to privacy in intimate decision-making, e.g., *Griswold v. Connecticut,* supra (striking down statute prohibiting use of contraceptives). What the *Griswold* Court found "repulsive to the notions of privacy surrounding the marriage relationship" was the prospect of "allow[ing] the police to search the sacred precincts of marital bedrooms for telltale signs of the use of contraceptives." Similarly, in *Lawrence v. Texas,* (2003), it was the criminalization of private sexual behavior that the Court found violative of the petitioners' liberty interest.

The marriage statute, which regulates only the act of obtaining a marriage license, does not intrude on any right that the plaintiffs have to privacy in their choices regarding procreation, an intimate partner or sexual relations. The plaintiffs' right to privacy in such matters does not require that the State officially endorse their choices in order for the right to be constitutionally vindicated.

Although some of the privacy cases also speak in terms of personal autonomy, no court has ever recognized such an open-ended right. Such decisions are protected not because they are important, intimate, and personal, but because the right or liberty at stake is "so deeply rooted in our history and traditions, or so fundamental to our concept of constitutionally ordered liberty" that it is protected by due process. Accordingly, the Supreme Court has concluded that while the decision to refuse unwanted medical treatment is fundamental because it is deeply rooted in our nation's history and tradition, the equally personal and profound decision to commit suicide is not because of the absence of such roots. While the institution of marriage is deeply rooted in the history and traditions of our country and our State, the right to marry someone of the same sex is not. No matter how personal or intimate a decision to marry someone of the same sex might be, the right to make it is not guaranteed by the right of personal autonomy. . . .

The consequences of deeming a right to be "fundamental" are profound, and this court, as well as the Supreme Court, has been very cautious in recognizing them. Such caution is required by separation of powers principles. If a right is found to be "fundamental," it is, to a great extent, removed from "the arena of public debate and legislative action"; utmost care must be taken when breaking new ground in this field "lest the liberty protected by the Due Process Clause be subtly transformed into the policy preferences of [judges]." Applying this limiting principle, the Supreme Court, as noted above, declined to recognize a fundamental right to physician-assisted suicide. While recognizing that public attitudes toward assisted suicide are currently the subject of "earnest and profound debate," the Court nevertheless left the continuation and resolution of that debate to the political arena, "as it should be in a democratic society." . . .

Although public attitudes toward marriage in general and same-sex marriage in particular have changed and are still evolving, "the asserted contemporary concept of marriage and societal interests for which [plaintiffs] contend" are "manifestly [less] deeply founded" than the "historic institution" of marriage. In such circumstances, the law with respect to same-sex marriages must be left to develop through legislative processes, subject to the constraints of rationality, lest the court be

viewed as using the liberty and due process clauses as vehicles merely to enforce its own views regarding better social policies, a role for which the court is particularly ill suited.

RATIONAL BASIS

The marriage statute satisfies the rational basis standard. The burden of demonstrating that a statute does not satisfy the rational basis standard rests on the plaintiffs. It is a weighty one. "[A] reviewing court will presume a statute's validity, and make all rational inferences in favor of it. . . . The Legislature is not required to justify its classifications, nor provide a record or finding in support of them." The "time tested wisdom of the separation of powers" requires courts to avoid "judicial legislation in the guise of new constructions to meet real or supposed new popular viewpoints, preserving always to the Legislature alone its proper prerogative of adjusting the statutes to changed conditions."

In analyzing whether a statute satisfies the rational basis standard, we look to the nature of the classification embodied in the enactment, then to whether the statute serves a legitimate State purpose, and finally to whether the classification is reasonably related to the furtherance of that purpose. . . .

STATE PURPOSE

The court's opinion concedes that the civil marriage statute serves legitimate State purposes, but further investigation and elaboration of those purposes is both helpful and necessary. Civil marriage is the institutional mechanism by which societies have sanctioned and recognized particular family structures, and the institution of marriage has existed as one of the fundamental organizing principles of human society.

Paramount among its many important functions, the institution of marriage has systematically provided for the regulation of heterosexual behavior, brought order to the resulting procreation, and ensured a stable family structure in which children will be reared, educated, and socialized. Admittedly, heterosexual intercourse, procreation, and child care are not necessarily conjoined . . . , but an orderly society requires some mechanism for coping with the fact that sexual intercourse commonly results in pregnancy and childbirth. The institution of marriage is that mechanism.

The institution of marriage provides the important legal and normative link between heterosexual intercourse and procreation on the one hand and family responsibilities on the other. The partners in a marriage are expected to engage in exclusive sexual relations, with children the probable result and paternity presumed. . . . The alternative, a society without the institution of marriage, in which heterosexual intercourse, procreation, and child care are largely disconnected processes, would be chaotic.

The marital family is also the foremost setting for the education and socialization of children. Children learn about the world and their place in it primarily from those who raise them, and those children eventually grow up to exert some influence, great or small, positive or negative, on society. The institution of marriage encourages parents to remain committed to each other and to their children as they grow, thereby encouraging a stable venue for the education and socialization of children.

It is undeniably true that dramatic historical shifts in our cultural, political, and economic landscape have altered some of our traditional notions about marriage. Nevertheless, the institution of marriage remains the principal weave of our social fabric. It is difficult to imagine a State purpose more important and legitimate than ensuring, promoting, and supporting an optimal social structure within which to bear and raise children. At the very least, the marriage statute continues to serve this important State purpose.

RATIONAL RELATIONSHIP

The question we must turn to next is whether the statute, construed as limiting marriage to couples of the opposite sex, remains a rational way to further that purpose. In considering whether such a rational basis exists, we defer to the decision-making

process of the Legislature, and must make deferential assumptions about the information that it might consider and on which it may rely.

We must assume that the Legislature (1) might conclude that the institution of civil marriage has successfully and continually provided this structure over several centuries; (2) might consider and credit studies that document negative consequences that too often follow children either born outside of marriage or raised in households lacking either a father or a mother figure, and scholarly commentary contending that children and families develop best when mothers and fathers are partners in their parenting; and (3) would be familiar with many recent studies that variously support the proposition that children raised in intact families headed by same-sex couples fare as well on many measures as children raised in similar families headed by opposite-sex couples support the proposition that children of same-sex couples fare worse on some measures; or reveal notable differences between the two groups of children that warrant further study. . . .

Taking all of this available information into account, the Legislature could rationally conclude that a family environment with married opposite-sex parents remains the optimal social structure in which to bear children, and that the raising of children by same-sex couples, who by definition cannot be the two sole biological parents of a child and cannot provide children with a parental authority figure of each gender, presents an alternative structure for child rearing that has not yet proved itself beyond reasonable scientific dispute to be as optimal as the biologically based marriage norm. Working from the assumption that a recognition of same-sex marriages will increase the number of children experiencing this alternative, the Legislature could conceivably conclude that declining to recognize same-sex marriages remains prudent until empirical questions about its impact on the upbringing of children are resolved.

The fact that the Commonwealth currently allows same-sex couples to adopt, does not affect the rationality of this conclusion. The eligibility of a child for adoption presupposes that at least one of the child's biological parents is unable or unwilling, for some reason, to participate in raising the child. In that sense, society has "lost" the optimal setting in which to raise that child—it is simply not available.

That the Legislature permits single-parent adoption does not mean that it has endorsed single parenthood as an optimal setting in which to raise children or views it as the equivalent of being raised by both of one's biological parents. The same holds true with respect to same-sex couples—the fact that they may adopt children means only that the Legislature has concluded that they may provide an acceptable setting in which to raise children who cannot be raised by both of their biological parents. The Legislature may rationally permit adoption by same-sex couples yet harbor reservations as to whether parenthood by same-sex couples should be affirmatively encouraged to the same extent as parenthood by the heterosexual couple whose union produced the child.

In addition, the Legislature could conclude that redefining the institution of marriage to permit same-sex couples to marry would impair the State's interest in promoting and supporting heterosexual marriage as the social institution that it has determined best normalizes, stabilizes, and links the acts of procreation and child rearing. . . .

CONCLUSION

While "[t]he Massachusetts Constitution protects matters of personal liberty against government incursion as zealously, and often more so, than does the Federal Constitution," this case is not about government intrusions into matters of personal liberty. It is not about the rights of same-sex couples to choose to live together, or to be intimate with each other, or to adopt and raise children together. It is about whether the State must endorse and support their choices by changing the institution of civil marriage to make its benefits, obligations, and responsibilities applicable to them. While the courageous efforts of many have resulted in increased dignity, rights, and respect for gay and lesbian members of our community, the issue presented here is a profound one, deeply rooted in social policy, that must, for now, be the subject of legislative, not judicial, action.

REVIEW AND DISCUSSION QUESTIONS

1. How does the majority respond to the argument that same-sex marriage should be permitted because the existing law does not explicitly rule it out?
2. Why does the majority think marriage is a fundamental right?
3. Explain and assess the majority's response to the three legislative rationales for the existing law.
4. In what way does the dissent argue that tradition is relevant to the question of whether marriage is a fundamental right that applies to homosexual couples?
5. Describe the rational basis standard and how it is applied to cases.
6. Why do the dissenters think that banning gay marriage passes the "rational basis" test?
7. Discuss whether this case was decided based on the morality of gay marriage, political issues involving separation of power and the role of the courts, or both.

Judicial Activism and Gay Marriage: A Debate

Lawrence Tribe and Richard Parker

Courts often make controversial decisions that restrict the power of elected officials by extending rights against government. One such case was *Goodridge v. Department of Public Health,* in which the Massachusetts Supreme Judicial Court held that the state constitution required that gays be allowed to marry. In this selection, two law professors—both of whom are sympathetic to same-sex marriage—debate whether the court overstepped its authority in demanding that the right to marry be extended to gays.

On the 50th anniversary of Brown v. Board of Education, *how do you predict* Goodridge *will be thought of in half a century?*

Professor Laurence Tribe: I would expect *Goodridge* at mid-century will look very much like the California Supreme Court's pathbreaking 1948 ruling on interracial marriage looks today: like a decision that prefigured things to come.

Professor Richard Parker: How *Goodridge* is thought of 50 years from now will depend on what happens in the meantime.

In 50 years, gays and lesbians may be (I hope) fully integrated into the American community. Same-sex marriage may be legal in every state. That future is eminently imaginable. It is in keeping with the narrative of one-way "progress" that we sometimes tell ourselves as if it were a law of history. If that is what happens, *Goodridge* may very well be celebrated as the crucial step between here and there. It will be mainly lawyers, though, who celebrate it as such.

We lawyers often "forget" that it is politics that drives and legitimates changes in the law, not vice versa. At bottom, even judge-made law is in and of politics. Its success is not a matter of hurling a lightning bolt from "above." To take root and effect it must have, or soon win, popular support. Thus it is—and should be—tethered to democratic politics. Political actors—whether a president, a legislature or a court—must draw on, and participate in, a typically gradual process of public persuasion. The 50-year struggle over "enforcement" of *Brown* illustrates the point. The same will hold true for *Goodridge* . . . in the event that it is, in fact, counted a "success" in 2054.

More than a realistic estimation of judicial influence is at stake here. The notion that court mandates, by themselves, "disentrench" the political status quo or "catalyze" political change isn't just a self-serving mistake. It nurtures a perverse idea that

it is the very mission of courts to act "heroically" by doing unpopular things—defying public opinion or coercing it. The idea is perverse because it is antidemocratic and because it tends to be self-defeating. Today, some who celebrate *Brown* misrepresent it by wrapping it in a story of judicial "heroism." They thus mislead those who are inspired by *Brown.* It may be that some such judicial triumphalism influenced the majority in *Goodridge*—and this misconception may turn out, in the end, to affect its fate.

One thing you know about politics is that you never know, for sure, what is going to happen. What if—as a matter of political preparation and timing—*Goodridge* "came too soon" or "went too far too fast"? What if it is widely experienced as imposing the values of a privileged, unelected minority on a majority? What if a powerful "backlash" movement develops, as the right-to-life movement developed after *Roe*? What if, in response, barriers to gay and lesbian equality are set up by state or federal legislation or constitutional amendments? Might not *Goodridge* be fingered, then, as the crucial step between here and there? If so, its significance might again be exaggerated. But no more so, surely, than in the other story, the story of court-driven "progress."

There's another dimension, then, on which *Goodridge* might be remembered. It might be remembered as emblematic of a period in which self-styled legal "progressives" lost touch with the lessons of legal realism and the values of majoritarian democracy. Its formalist rhetoric is at odds with the open, controvertible argument of *Brown.* The peremptoriness of its claim to find something radically new in generic abstractions of a state constitution without much attention to anything in particular about that constitution or that state—along with the sloppiness, even the cynicism, of its supposedly value-neutral "rational basis" standard of review—looks disingenuous enough now. How will the opinion look once the drama of the moment has passed? Its tone of righteous command-and-control and its disrespect for electoral and legislative politics—its too easy willingness to attribute irrationality and prejudice to people embracing a traditional idea of marriage—are startling enough today. Might it seem, in 50 years, as arrogant as the antipopular prejudice of an 18th-century aristocrat?

It's not too soon, now, to turn a cool eye on *Goodridge*, however much we believe in the justice of its mandate. Even those who assume an inevitable Onward March of Equality ought to think about it twice, so long as the March, in their mind, includes political equality—which is to say, one person, one vote; which is to say, majority rule.

In 1948 [in Perez v. Sharp*], the California State Supreme Court was the first state high court to declare a ban on interracial marriage unconstitutional, despite strong public opposition. How does the* Goodridge *decision compare?*

Tribe: To 14-year-old Emmett Till, a black boy lynched and murdered in Mississippi in 1955 because he had dared to say "Bye-bye, baby" to a white woman, and to 21-year-old Matthew Shepard, a gay white boy robbed, beaten senseless, and then tied to a fence and left to die in Wyoming in 1998 for being gay, the differences between the history and structure of white supremacy from chattel slavery to the Black Codes to Jim Crow, and the structure and history of homophobia, gay-baiting and gay-bashing weren't awfully important. At bottom, both were the innocent victims of a bigotry born of fear, hammered into hate, ignited by rage and, in the end, enacted as death.

When I've compared the ban on interracial marriage that the U.S. Supreme Court didn't strike down until 1967 [in *Loving v. United States*] with the ban on same-sex marriage, I've sometimes been told that I'm mixing apples and oranges because the laws against so-called miscegenation reflect the view that persons of African-American descent are inferior and would pollute the proud genetic heritage of the Aryan race, whereas the refusal to recognize same-sex unions as forms of marriage need cast no aspersions on gays when it merely reflects the basic nature of things. We've all heard the mantra: "It was Adam and Eve in the Garden of Eden, not Adam and Steve."

Of course, from Steve's perspective, being told that it's not enough for him and Adam to love one another and to commit to one another until death do them part because Steve is a gay male rather than a straight female (and that, in such circumstances, Steve and Adam will just have to be content—at best—with something called a "civil

union") must feel like Eve's being told she can't marry Adam either because she's white and Adam is black—nothing against her as a white woman, you understand; after all, the rule is no less a restriction on the liberty of black men.

Some, ignorant of history, say that, when California's highest court saw through the phony symmetry of the ban on interracial marriage in 1948, a great majority of the nation's electorate had already come to the same conclusion, so the court wasn't making new law so much as it was ratifying an existing consensus. But the truth is that, as late as 1958, polls reported more than 95 percent of whites still disapproving of marriages between blacks and whites. If anything, therefore, the SJC hasn't ventured as far beyond the consensus as its California precursor had.

Nor is there much to the argument that the mixed-sex character of the marital unit simply "feels" more deeply woven into the "marriage" concept itself than could ever have been said about the same-race character of a marital unit. In December 1912, when Representative Seaborn Roddenbery of Georgia proposed a constitutional amendment to uphold the sanctity of marriage—an amendment that would have said "Intermarriage between Negroes or persons of color and Caucasians . . . is forever prohibited"—he spoke in terms that should make proponents of a federal marriage amendment, designed to overturn decisions like *Goodridge* and to prevent their ilk from seeing the light of day again either under any state constitution or under the U.S. Constitution, feel right at home. The Roddenbery Amendment was said to be necessary to "exterminate now this debasing, ultrademoralizing, un-American and inhuman leprosy." Such "marriages" were that in name only to Mr. Roddenbery. They were, in truth, "abhorrent and repugnant" counterfeits of the "sacred marital estate."

I don't want to overstate the parallel. No doubt some who have nothing against gays or lesbians and who wouldn't mind learning that their sons were gay or their daughters were lesbians—people who don't want to force anybody back into the closet and who'd be happy to accord same-sex couples civil unions possessing every legal incident of marriage—simply balk at extending the "M" word to a relationship they think is neither better nor worse than marriage, simply different. But the issue isn't resolved by good intentions, which we all know can pave one hell of a nasty road. Decisive is the fact that, intended or not, the signal sent to the same-sex couples who are relegated to the "separate but equal" status of civil union is that their relationship is at best a facsimile of the real thing: a good copy, maybe, but ultimately fake. Using a state's legal system to deliver such a signal of second-class status is inconsistent with the most basic axioms of equal citizenship.

Parker: Again, this question suggests judicial-triumphalism. Being "the first" to do something which is analogized to something "first" done by a state court combating racial oppression, something that ultimately won out to become the law of the nation, certainly sounds good. In the *Goodridge* opinion itself, the analogy to antimiscegenation laws is waved like a flag to be followed into battle. What if we look at it as an argument rather than a flag? Shouldn't we consider whether to distinguish a supposed precedent before jumping to follow it in a very different context?

What I have in mind is not a comparison, on the facts, between degrees of oppression of African-Americans, on one hand, and of gays and lesbians, on the other. Nor is it a comparison between the significance of barriers to marriage in the two contexts. Though there are differences along these lines, and though none were addressed by the *Goodridge* court, what I have in mind, instead, is a comparison between the decision in *Goodridge* and the California decision in *Perez*—and ultimately the U.S. Supreme Court decision in *Loving*—as political decisions, political actions.

This kind of comparison is the most difficult. The reason is that politics is more complicated than logic or even social history. It calls for context-sensitive judgments of strategy and tactics, an assessment of preparation and timing, of what resistance may be anticipated and overcome, what popular support gained, by what actions and arguments, and when. I don't know enough about California in 1948 to make a satisfactory comparison. But I know a little, and I know the standard against which to make the comparison: Was the law that the court attacked already politically weak enough to allow for judicial success? Was

public opinion ready to entertain, then embrace, the judicial mandate? Courts, I think, should push only on doors that are ajar, that the people are already prepared to pass through—not on doors that, for the moment, are firmly shut, much less ones that may well be locked in response to an unexpected judicial push.

Consider the situation facing the California court. In 1948, interracial marriage was already permitted in more than one-third of the states. The decision to permit it had been made by legislatures, not courts. Among the remaining states, California's antimiscegenation law was unique—uniquely weak. It attached no criminal penalty to interracial marriage, and according to a historian, it was "already largely eviscerated in practice." Indeed, one justice concurring in *Perez* went so far as to invoke the "attitude of comparative indifference [to miscegenation] on the part of the [California] Legislature" and "the absence of any clearly expressed public sentiment or policy [about it]" along with "strong indications of legislative intentions which point the [other] way."

The U.S. Supreme Court, then, waited two decades to confront the issue. It, too, pursued a path of political caution and respect for majority rule. It, too, did not defy democratic politics. It waited to follow it. Between 1948 and 1967—while Congress moved to prohibit racial discrimination—many state legislatures repealed antimiscegenation laws. By the time *Loving v. U.S.* [1967] was finally decided, interracial marriage was already permitted in two-thirds of the states.

Now, contrast *Goodridge.* No state permitted same-sex marriage—or ever had. There was no history of non-enforcement of the traditional limitation on marriage. Two courts, in Hawaii and Alaska, pushed at the door, tentatively, in the last decade. These initiatives were quickly repulsed by the voters of each state—locking in new barriers to same-sex marriage—and then resisted by further state and federal legislation. Was there reason to believe the people of Massachusetts were more prepared, in 2003, to embrace same-sex marriage than the people of Hawaii a few years before? I doubt it.

What's more, the Massachusetts court lacked a vital resource of constitutional argument that had been available to the California court. In 1948, suspicion of racial discrimination was well-established in American constitutional law. The *Perez* opinion invoked it. There is, to this day, no well-established suspicion of discrimination by sexual orientation. *Goodridge* was left to put its emphasis on the importance of marriage—which played into the hands of its critics whose condemnation of the decision proceeded from just the same premise.

There was available, of course, an alternative. The Supreme Court of Vermont had taken an incremental step, requiring the institution of civil unions. That ignited a political struggle but seems to have stuck. You might think the Massachusetts court would have opted to follow Vermont. (Even the Netherlands, the first country to legalize same-sex marriage, had begun by instituting civil unions.) But no. What the *Goodridge* court lacked in political warrant for its decision, it made up for in self-righteous self-assertion. The risk that seemed to concern it most was the risk that some other court, later on down the road, would be "first."

What do you say to those who argue that Goodridge *could open the door to legalizing other unions, such as those between siblings or a triad?*

Tribe: When confronted with such hypotheticals, we need to ask ourselves how plausible is it to imagine the dynamic generated by discrimination against same-sex couples—a dynamic constituted by violent intolerance toward those open about their intimate relations and by equally devastating self-erasure by those closeted about their sexual orientations—at work in these other, very different, contexts. Incest laws draw circles around individuals, defining the finite set of family members so closely tied by blood or adoption that sexual intimacy becomes too dangerous or volatile for society to sanction. These restrictions no doubt inflict a heavy burden on particular hapless individuals whose misfortune it is to lust after or to fall in love with a family member, but such tightly drawn circles bear no real resemblance to the broad lines cutting oppressively across society to rule half the adult population off limits as sexual or marital partners for a distinct and despised minority. So, too, the circles that our adultery and bigamy laws have drawn around married couples

have established partitions that fall with an undeniably cruel weight upon individuals who fall in love or lust with someone else's spouse. But these laws likewise cut no wide swath through the population to limit the options open to any particular oppressed minority.

Parker: Since the life of the law is (and should be) democratic politics—not logic—the legalization of other unions will occur only if and when there is widespread popular sentiment favoring such an expansion of marriage. I don't foresee that soon, do you? . . .

REVIEW AND DISCUSSION QUESTIONS

1. Explain the argument that Tribe makes in defense of the rights of gays to marry.
2. How does Parker respond to Tribe's position? Is his position based only on the potential consequences of the Court's ruling, or does he also think there is an important matter of political principle at stake?
3. Will the *Goodridge* decision open the door to polygamous or incestuous marriages?
4. Is the *Goodridge* decision required by the precedents of *Perez v. Sharp* and *Loving v. United States*?

CHAPTER 21

Freedom of Speech

This section looks at freedom of speech. Enshrined in the First Amendment, it is one of our most cherished political and constitutional values, but it has often posed difficult legal conundrums for the courts and embroiled them in political controversies. Some of these legal battles, for example, over flag burning, advocacy of Nazism, obscenity, and the First Amendment rights of students, are reflected in the cases that follow. But first we begin with John Stuart Mill's reasoned and articulate defense of liberty of thought and discussion.

Of the Liberty of Thought and Discussion

John Stuart Mill

In the following selection, taken from his classic work *On Liberty,* John Stuart Mill provides what may be the best known and most widely quoted defense of freedom of conscience and speech ever written. Arguing that protecting freedom of conscience and speech is of vital social importance, Mill vigorously upholds freedom of opinion, regardless of whether the beliefs in question are true or false. Indeed, he argues, it is especially important that society permit false opinions to be freely expressed.

The time, it is to be hoped, is gone by, when any defense would be necessary of the "liberty of the press" as one of the securities against corrupt or tyrannical government. No argument, we may suppose, can now be needed against permitting a legislature or an executive, not identified in interest with the people, to prescribe opinions to them, and determine what doctrines or what arguments they shall be allowed to hear. . . . Let us suppose . . . that government is entirely at one with the people, and never thinks of exerting any power of coercion unless in agreement with what it conceives to be their voice. But I deny the right of the people to exercise such coercion, either by themselves or by their government. The power itself is illegitimate. The best government has no more title to it than the worst. It is as noxious, or more noxious, when exerted in accordance with public opinion than when in opposition to it. If all mankind minus one were of one opinion, mankind would be no more

justified in silencing that one person than he, if he had the power, would be justified in silencing mankind. Were an opinion a personal possession of no value except to the owner, if to be obstructed in the enjoyment of it were simply a private injury, it would make some difference whether the injury was inflicted only on a few persons or on many. But the peculiar evil of silencing the expression of an opinion is that it is robbing the human race, posterity as well as the existing generation—those who dissent from the opinion, still more than those who hold it. If the opinion is right, they are deprived of the opportunity of exchanging error for truth; if wrong, they lose, what is almost as great a benefit, the clearer perception and livelier impression of truth produced by its collision with error.

It is necessary to consider separately these two hypotheses, each of which has a distinct branch of the argument corresponding to it. We can never be sure that the opinion we are endeavoring to stifle is a false opinion; and if we were sure, stifling it would be an evil still.

First, the opinion which it is attempted to suppress by authority may possibly be true. Those who desire to suppress it, of course, deny its truth; but they are not infallible. They have no authority to decide the question for all mankind and exclude every other person from the means of judging. To refuse a hearing to an opinion because they are sure that it is false to assume that *their* certainty is the same thing as *absolute* certainty. All silencing of discussion is an assumption of infallibility. Its condemnation may be allowed to rest on this common argument, not the worse for being common.

Unfortunately for the good sense of mankind, the fact of their fallibility is far from carrying the weight in their practical judgment which is always allowed to it in theory; for while everyone well knows himself to be fallible, few think it necessary to take any precautions against their own fallibility, or admit the supposition that any opinion of which they feel very certain may be one of the examples of the error to which they acknowledge themselves to be liable. . . .

The objection likely to be made to this argument would probably take some such form as the following. There is no greater assumption of infallibility in forbidding the propagation of error than in any other thing which is done by public authority on its own judgment and responsibility. . . . It is the duty of governments, and of individuals, to form the truest opinions they can; to form them carefully, and never impose them upon others unless they are quite sure of being right. But when they are sure (such reasoners may say), it is not conscientiousness but cowardice to shrink from acting on their opinions and allow doctrines which they honestly think dangerous to the welfare of mankind, either in this life or in another, to be scattered abroad without restraint, because other people, in less enlightened times, have persecuted opinions now believed to be true. . . . There is no such thing as absolute certainty, but there is assurance sufficient for the purposes of human life. We may, and must, assume our opinion to be true for the guidance of our own conduct; and it is assuming no more when we forbid bad men to pervert society by the propagation of opinions which we regard as false and pernicious.

I answer, that it is assuming very much more. There is the greatest difference between presuming an opinion to be true because, with every opportunity for contesting it, it has not been refuted, and assuming its truth for the purpose of not permitting its refutation. Complete liberty of contradicting and disproving our opinion is the very condition which justifies us in assuming its truth for purposes of action; and on no other terms can a being with human faculties have any rational assurance of being right. . . .

In the present age—which has been described as "destitute of faith, but terrified at skepticism"—in which people feel sure, not so much that their opinions are true as that they should not know what to do without them—the claims of an opinion to be protected from public attack are rested not so much on its truth as on its importance to society. There are, it is alleged, certain beliefs so useful, not to say indispensable, to well-being that it is as much the duty of governments to uphold those beliefs as to protect any other of the interests of society. . . . The usefulness of an opinion is itself [a] matter of opinion—as disputable, as open to discussion, and requiring discussion as much as the opinion itself. There is the same need of an infallible judge of opinions to decide an opinion to be noxious as to decide it to be false, unless the opinion condemned has full opportunity of defending itself. And it will

not do to say that the heretic may be allowed to maintain the utility or harmlessness of his opinion, though forbidden to maintain its truth. The truth of an opinion is part of its utility. If we would know whether or not it is desirable that a proposition should be believed, is it possible to exclude the consideration of whether or not it is true? In the opinion, not of bad men, but of the best men, no belief which is contrary to truth can be really useful. . . .

The dictum that truth always triumphs over persecution is one of those pleasant falsehoods which men repeat after one another till they pass into commonplaces, but which all experience refutes. History teems with instances of truth put down by persecution. If not suppressed forever, it may be thrown back for centuries. To speak only of religious opinions: the Reformation broke out at least twenty times before Luther, and was put down. . . . Even after the era of Luther, wherever persecution was persisted in, it was successful. . . .

Let us now pass to the second division of the argument, and dismissing the supposition that any of the received opinions may be false, let us assume them to be true and examine into the worth of the manner in which they are likely to be held when their truth is not freely and openly canvassed. However unwillingly a person who has a strong opinion may admit the possibility that his opinion may be false, he ought to be moved by the consideration that, however true it may be, if it is not fully, frequently, and fearlessly discussed, it will be held as a dead dogma, not a living truth. . . .

If the cultivation of the understanding consists in one thing more than in another, it is surely in learning the grounds of one's own opinions. Whatever people believe, on subjects on which it is of the first importance to believe rightly, they ought to be able to defend against at least the common objections. . . . On every subject on which difference of opinion is possible, the truth depends on a balance to be struck between two sets of conflicting reasons. Even in natural philosophy, there is always some other explanation possible of the same facts; some geocentric theory instead of heliocentric, some phlogiston instead of oxygen; and it has to be shown why that other theory cannot be the true one; and until this is shown, and until we know how it is shown, we do not understand the grounds of our opinion. But when we turn to subjects infinitely more complicated, to morals, religion, politics, social relations, and the business of life, three-fourths of the arguments for every disputed opinion consist in dispelling the appearances which favor some opinion different from it. . . . Nor is it enough that he should hear the arguments of adversaries from his own teachers, presented as they state them, and accompanied by what they offer as refutations. That is not the way to do justice to the arguments or bring them into real contact with his own mind. He must be able to hear them from persons who actually believe them, who defend them in earnest and do their very utmost for them. He must know them in their most plausible and persuasive form; he must feel the whole force of the difficulty which the true view of the subject has to encounter and dispose of, else he will never really possess himself of the portion of truth which meets and removes that difficulty. . . .

The fact . . . is that not only the grounds of the opinion are forgotten in the absence of discussion, but too often the meaning of the opinion itself. The words which convey it cease to suggest ideas, or suggest only a small portion of those they were originally employed to communicate. Instead of a vivid conception and a living belief, there remain only a few phrases retained by rote; or, if any part, the shell and husk only of the meanings is retained, the finer essence being lost. The great chapter in human history which this fact occupies and fills cannot be too earnestly studied and meditated on.

It is illustrated in the experience of almost all ethical doctrines and religious creeds. They are all full of meaning and vitality to those who originate them, and to the direct disciples of the originators. Their meaning continues to be felt in undiminished strength, and is perhaps brought out into even fuller consciousness, so long as the struggle lasts to give the doctrine or creed an ascendancy over other creeds. . . .

We have hitherto considered only two possibilities: that the received opinion may be false, and some other opinion, consequently, true; or that, the received opinion being true, a conflict with the opposite error is essential to a clear apprehension and deep feeling of its truth. But there is a commoner case than either of these: when the conflicting doctrines, instead of being one true and the other false, shared the truth between them, and the

nonconforming opinion is needed to supply the remainder of the truth of which the received doctrine embodies only a part. Popular opinions, on subjects not palpable to sense, are often true, but seldom or never the whole truth. They are a part of the truth, sometimes a greater, sometimes a smaller part, but exaggerated, distorted, and disjointed from the truths by which they ought to be accompanied and limited. Heretical opinions, on the other hand, are generally some of these suppressed and neglected truths, bursting the bonds which kept them down, and either seeking reconciliation with the truth contained in the common opinion, or fronting it as enemies, and setting themselves up, with similar exclusiveness, as the whole truth. The latter case is hitherto the most frequent, as, in the human mind, one-sidedness has always been the rule, and many-sidedness the exception. Hence, even in revolutions of opinion, one part of the truth usually sets while another rises. Even progress, which ought to superadd, for the most part only substitutes one partial and incomplete truth for another; improvement consisting chiefly in this, that the new fragment of truth is more wanted, more adapted to the needs of the time than that which it displaces. Such being the partial character of prevailing opinions, even when resting on a true foundation, every opinion which embodies somewhat of the portion of truth which the common opinion omits ought to be considered precious, with whatever amount of error and confusion that truth may be blended. . . .

We have now recognized the necessity to the mental well-being of mankind (on which all their other well-being depends) of freedom of opinion, and freedom of the expression of opinion, on four distinct grounds, which we will now briefly recapitulate:

First, if any opinion is compelled to silence, that opinion may, for aught we can certainly know, be true. To deny this is to assume our own infallibility.

Secondly, although the silenced opinion be an error, it may, and very commonly does, contain a portion of truth; and since the general or prevailing opinion on any subjects is rarely or never the whole truth, it is only by the collision of adverse opinions that the remainder of the truth has any chance of being supplied.

Thirdly, even if the received opinion be not only true, but the whole truth; unless it is suffered to be, and actually is, vigorously and earnestly contested, it will, by most of those who receive it, be held in the manner of a prejudice, with little comprehension or feeling of its rational grounds. And not only this, but fourthly, the meaning of the doctrine itself will be in danger of being lost or enfeebled, and deprived of its vital effect on the character and conduct: the dogma becoming a mere formal profession, inefficacious for good, but cumbering the ground and preventing the growth of any real and heartfelt conviction from reason or personal experience. . . .

Before quitting the subject of freedom of opinion, it is fit to take some notice of those who say that the free expression of all opinions should be permitted on condition that the manner be temperate, and do not pass the bounds of fair discussion. Much might be said on the impossibility of fixing where these supposed bounds are to be placed; for if the test be offense to those whose opinions are attacked, I think experience testifies that this offense is given whenever the attack is telling and powerful, and that every opponent who pushes them hard, and whom they find it difficult to answer, appears to them, if he shows any strong feeling on the subject, an intemperate opponent. . . . With regard to what is commonly meant by intemperate discussion, namely invective, sarcasm, personality, and the like, the denunciation of these weapons would deserve more sympathy if it were ever proposed to interdict them equally to both sides; but it is only desired to restrain the employment of them against the prevailing opinion; against the unprevailing they may not only be used without general disapproval, but will be likely to obtain for him who uses them the praise of honest zeal and righteous indignation. Yet whatever mischief arises from their use is greatest when they are employed against the comparatively defenseless; and whatever unfair advantage can be derived by any opinion from this mode of asserting it accrues almost exclusively to received opinions. . . . In general, opinions contrary to these commonly received can only obtain a hearing by studied moderation of language, and the most cautious avoidance of unnecessary offense, from which they hardly ever deviate even in a slight degree without losing ground; while unmeasured vituperation employed on the side of

the prevailing opinion really does deter people from professing contrary opinions, and from listening to those who profess them. For the interest, therefore, of truth and justice, it is far more important to restrain this employment of vituperative language than the other; and, for example, if it were necessary to choose, there would be much more need to discourage offensive attacks on infidelity than on religion. It is, however, obvious that law and authority have no business restraining either, while opinion ought, in every instance, to determine its verdict by the circumstances of the individual case; condemning everyone, on whichever side of the argument he places himself, in whose mode of advocacy either want of candor, or malignity, bigotry, or intolerance of feeling manifest themselves; but not inferring these vices from the side which a person takes, though it be the contrary side of the question of our own; and giving merited honor to everyone, whatever opinion he may hold, who has calmness to see the honesty to state what his opponents and their opinions really are, exaggerating nothing to their discredit, keeping nothing back which tells or can be supposed to tell, in their favor. This is the real morality of public discussion; and if often violated, I am happy to think that there are many controversialists who to a great extent observe it, and a still greater number who conscientiously strive towards it.

REVIEW AND DISCUSSION QUESTIONS

1. What values or purposes does Mill think are served by freedom of speech?
2. Why does Mill think that false speech should be tolerated or even encouraged?
3. Explain why Mill thinks those who would censor speech are assuming, falsely, that they are "infallible."
4. Does Mill think that truth, rationality, and free speech are related? Explain why or why not.
5. How does Mill respond to those who say that "offensive" speech may be banned?
6. Describe how Mill's discussion of free speech fits into his larger discussion of liberty and individuality in *On Liberty*.
7. Suppose somebody defends censorship of Nazi propaganda based on the claim that unless their ideas are suppressed Nazis might eventually gain power, as occurred in post–World War I Germany. How would Mill respond to such an argument? Is that response sound?

Flag Burning

Texas v. Johnson

After burning the U.S. flag as an act of political protest, Gregory Lee Johnson was convicted of desecrating a flag in violation of Texas law. The state of Texas, after losing in lower courts, appealed to the U.S. Supreme Court, which had to decide whether Johnson's conviction was consistent with the First Amendment's protection of freedom of speech. By a narrow 5 to 4 vote, the Court held that the Texas law was unconstitutional. Delivering the opinion of the Court, Justice Brennan argues that the state cannot "prescribe what shall be orthodox" by punishing symbolic actions such as flag burning. The way to preserve the flag's special role in our national life, he argues, is not to punish those who feel differently about this symbol but to persuade them that they are wrong. In their separate dissents, Justice Rehnquist and Justice Stevens reject the idea that the flag is just another symbol, toward which it would be unconstitutional to require minimal respect.

---

57 L. W. 4770 (1989).

Justice Brennan: As in *Spence* [*v. Washington*, a 1974 case on expressive conduct], "[w]e are confronted with a case of prosecution for the expression of an idea through activity," and "[a]ccordingly, we must examine with particular care the interests advanced by [petitioner] to support its prosecution." . . . Johnson was not, we add, prosecuted for the expression of just any idea; he was prosecuted for his expression of dissatisfaction with the policies of this country, expression situated at the core of our First Amendment values.

Moreover, Johnson was prosecuted because he knew that his politically charged expression would cause "serious offense." If he had burned the flag as a means of disposing of it because it was dirty or torn, he would not have been convicted of flag desecration under this Texas law: federal law designates burning as the preferred means of disposing of a flag "when it is in such condition that it is no longer a fitting emblem for display." . . .

If we are to hold that a state may forbid flag-burning wherever it is likely to endanger the flag's symbolic role, but allow it wherever burning a flag promotes that role—as where, for example, a person ceremoniously burns a dirty flag—we would be saying that when it comes to impairing the flag's physical integrity, the flag itself may be used as a symbol—as a substitute for the written or spoken word or a "short cut from mind to mind"—only in one direction. We would be permitting a state to "prescribe what shall be orthodox" by saying that one may burn the flag to convey one's attitude toward it and its referents only if one does not endanger the flag's representation of nationhood and national unity.

We never before have held that the government may ensure that a symbol be used to express only one view of that symbol or its referents. . . .

We are fortified in today's conclusion by our conviction that forbidding criminal punishment for conduct such as Johnson's will not endanger the special role played by our flag or the feelings it inspires. To paraphrase Justice [Oliver Wendell] Holmes, we submit that nobody can suppose that this one gesture of an unknown man will change our nation's attitude towards its flag. . . . Indeed, Texas's argument that the burning of an American flag "is an act having a high likelihood to cause a breach of peace," . . . and its statute's implicit assumption that physical mistreatment of the flag will lead to "serious offense," tend to confirm that the flag's special role is not in danger; if it were, no one would riot or take offense because a flag had been burned.

We are tempted to say, in fact, that the flag's deservedly cherished place in our community will be strengthened, not weakened, by our holding today. Our decision is a reaffirmation of the principles of freedom and inclusiveness that the flag best reflects, and of the conviction that our toleration of criticism such as Johnson's is a sign and source of our strength. Indeed, one of the proudest images of our flag, the one immortalized in our own national anthem, is of the bombardment it survived at Fort McHenry. It is the nation's resilience, not its rigidity, that Texas sees reflected in the flag—and it is that resilience that we reassert today.

The way to preserve the flag's special role is not to punish those who feel differently about these matters. It is to persuade them that they are wrong. "To courageous, self-reliant men, with confidence in the power of free and fearless reasoning applied through the processes of popular government, no danger flowing from speech can be deemed clear and present, unless the incidence of the evil apprehended is so imminent that it may befall before there is opportunity for full discussion. If there be time to expose through discussion the falsehood and fallacies, to avert the evil by the processes of education, the remedy to be applied is more speech, not enforced silence." . . . And, precisely because it is our flag that is involved, one's response to the flag-burner may exploit the uniquely persuasive power of the flag itself. We can imagine no more appropriate response to burning a flag than waving one's own, no better way to counter a flag-burner's message than by saluting the flag that burns, no surer means of preserving the dignity even of the flag that burned than by—as one witness here did—according its remains a respectful burial. We do not consecrate the flag by punishing its desecration, for in doing so we dilute the freedom that this cherished emblem represents.

Johnson was convicted for engaging in expressive conduct. The state's interest in preventing breaches of the peace does not support his conviction because Johnson's conduct did not threaten to disturb the peace. Nor does the state's interest in

preserving the flag as a symbol of nationhood and national unity justify his criminal conviction for engaging in political expression. The judgment of the Texas Court of Criminal Appeals is therefore affirmed.

Justice Rehnquist, dissenting: In holding this Texas statute unconstitutional, the court ignores Justice Holmes's familiar aphorism that "a page of history is worth a volume of logic." . . . For more than 200 years, the American flag has occupied a unique position as the symbol of our nation, a uniqueness that justifies a governmental prohibition against flag burning in the way respondent Johnson did here. . . .

In the First and Second World Wars, thousands of our countrymen died on foreign soil fighting for the American cause. At Iwo Jima in the Second World War, United States Marines fought hand-to-hand against thousands of Japanese. By the time the Marines reached the top of Mount Suribachi, they raised a piece of pipe upright and from one end fluttered a flag. That ascent had cost nearly 6,000 American lives. . . .

During the Korean War, the successful amphibious landing of American troops at Inchon was marked by the raising of an American flag within an hour of the event. . . .

The government is simply recognizing as a fact the profound regard for the American flag created by that history when it enacts statutes prohibiting the disrespectful public burning of the flag.

The court concludes its opinion with a regrettably patronizing civics lecture, presumably addressed to members of both houses of Congress, the members of the 48 state legislatures that enacted prohibitions against flag burning and the troops fighting under that flag in Vietnam who objected to its being burned: "The way to preserve the flag's special role is not to punish those who feel differently about these matters. It is to persuade them that they are wrong." . . .

The court's role as the final expositor of the Constitution is well established, but its role as a platonic guardian admonishing those responsible to public opinion as if they were truant school children has no similar place in our system of government. The cry of "no taxation without representation" animated those who revolted against the English crown to found our nation—the idea that those who submitted to government should have some say as to what kind of laws would be passed. Surely one of the high purposes of a democratic society is to legislate against conduct that is regarded as evil and profoundly offensive to the majority of people—whether it be murder, embezzlement, pollution or flag burning.

Our Constitution wisely places limits on powers of legislative majorities to act, but the declaration of such limits by this court "is, at all times, a question of much delicacy, which ought seldom, if ever, to be decided in the affirmative, in a doubtful case." . . . Uncritical extension of constitutional protection to the burning of the flag risks the frustration of the very purpose for which organized governments are instituted. The court decides that the American flag is just another symbol, about which not only must opinions pro and con be tolerated, but for which the most minimal public respect may not be enjoined. The government may conscript men into the armed forces where they must fight and perhaps die for the flag, but the government may not prohibit the public burning of the banner under which they fight. I would uphold the Texas statute as applied in this case.

Justice Stevens, dissenting: As the court analyzes this case, it presents the question whether the state of Texas, or indeed the federal government, has the power to prohibit the public desecration of the American flag. The question is unique. In my judgment, rules that apply to a host of other symbols, such as state flags, armbands or various privately promoted emblems of political or commercial identity, are not necessarily controlling. Even if flag burning could be considered just another species of symbolic speech under the logical application of the rules that the court has developed in its interpretation of the First Amendment in other contexts, this case has an intangible dimension that makes those rules inapplicable.

A country's flag is a symbol of more than "nationhood and national unity." . . . It also signifies the ideas that characterize the society that has chosen that emblem as well as the special history that has animated the growth and power of those ideas.

The fleurs-de-lis and the tricolor both symbolized "nationhood and national unity," but they had vastly different meanings. The message conveyed by some flags—the swastika, for example—may survive long after it has outlived its usefulness as a symbol of regimented unity in a particular nation.

So it is with the American flag. It is more than a proud symbol of the courage, the determination and the gifts of nature that transformed 13 fledgling colonies into a world power. It is a symbol of freedom, of equal opportunity, of religious tolerance and of goodwill for other peoples who share our aspirations. The symbol carries its message to dissidents both at home and abroad who may have no interest at all in our national unity or survival.

The value of the flag as a symbol cannot be measured. Even so, I have no doubt that the interest in preserving that value for the future is both significant and legitimate. Conceivably that value will be enhanced by the court's conclusion that our national commitment to free expression is so strong that even the United States as ultimate guarantor of that freedom is without power to prohibit the desecration of its unique symbol. But I am unpersuaded. . . .

The case has nothing to do with "disagreeable ideas." . . . it involves disagreeable conduct that, in my opinion, diminishes the value of an important national asset.

The court is therefore quite wrong in blandly asserting that respondent "was prosecuted for his expression of dissatisfaction with the policies of this country, expression situated at the core of our First Amendment values." . . . Respondent was prosecuted because of the method he chose to express his dissatisfaction with those policies. Had he chosen to spray paint—or perhaps convey with a motion picture projector—his message of dissatisfaction on the facade of the Lincoln Memorial, there would be no question about the power of the government to prohibit his means of expression. The prohibition would be supported by the legitimate interest in preserving the quality of an important national asset. Though the asset at stake in this case is intangible, given its unique value, the same interest supports a prohibition on the desecration of the American flag.

The ideas of liberty and equality have been an irresistible force in motivating leaders like Patrick Henry, Susan B. Anthony and Abraham Lincoln, schoolteachers like Nathan Hale and Booker T. Washington, the Philippine Scouts who fought at Bataan, and the soldiers who scaled the bluff at Omaha Beach. If those ideas are worth fighting for—and our history demonstrates that they are—it cannot be true that the flag that uniquely symbolizes their power is not itself worthy of protection from unnecessary desecration.

I respectfully dissent.

REVIEW AND DISCUSSION QUESTIONS

1. What does Texas argue is the danger in allowing flag burning?
2. What is the "fixed star" in our constitutional system that Justice Brennan argues is at the heart of this issue?
3. Why, according to Justice Brennan, might allowing such political protests actually *increase* people's sense of patriotism?
4. Why does Justice Rehnquist think government is justified in prohibiting flag burning?
5. On what basis does Justice Stevens dissent?
6. How might it be argued that preventing flag burning is *not* tantamount to establishing political orthodoxy?

Nazi Marches

Village of Skokie v. National Socialist Party

Skokie, Illinois, was the home of more than forty thousand Jews and five to seven thousand survivors of Nazi concentration camps. When the National Socialist Party (the American Nazi Party) tried to march in Skokie, the village won an injunction placing various restrictions on the march. An appeals court modified that injunction but allowed a ban on displaying the swastika to stand. Here the Supreme Court of Illinois considers an appeal by the Nazi leader, Frank Collin, of the lower court's ban. (The U.S. Supreme Court later refused to reconsider this decision of the Illinois Supreme Court.)

Per Curiam: [D]efendant Frank Collin, who testified that he was "party leader," stated that on or about March 20, 1977, he sent officials of the plaintiff village a letter stating that the party members and supporters would hold a peaceable, public assembly in the village on May 1, 1977, to protest the Skokie Park District's requirement that the party procure $350,000 of insurance prior to the party's use of the Skokie public parks for public assemblies. The demonstration was to begin at 3 p.m., last 20 to 30 minutes, and consist of 30 to 50 demonstrators marching in single file, back and forth, in front of the village hall. The marchers were to wear uniforms which include a swastika emblem or armband. They were to carry a party banner containing a swastika emblem and signs containing such statements as "White Free Speech," "Free Speech for the White Man," and "Free Speech for White America." The demonstrators would not distribute handbills, make any derogatory statements directed to any ethnic or religious group, or obstruct traffic. They would co-operate with any reasonable police instructions or requests.

At the hearing on plaintiff's motion for an "emergency injunction" a resident of Skokie testified that he was a survivor of the Nazi holocaust. He further testified that the Jewish community in and around Skokie feels the purpose of the march in the "heart of the Jewish population" is to remind the two million survivors "that we are not through with you" and to show "that the Nazi threat is not over, it can happen again." Another resident of Skokie testified that as the result of defendants' announced intention to march in Skokie, 15 to 18 Jewish organizations, within the village and surrounding area, were called and a counterdemonstration of an estimated 12,000 to 15,000 people was scheduled for the same day. There was opinion evidence that defendants' planned demonstration in Skokie would result in violence. . . .

In defining the constitutional rights of the parties who come before this court, we are, of course, bound by the pronouncements of the United States Supreme Court in its interpretation of the United States Constitution. The decisions of that court, particularly *Cohen v. California* . . . in our opinion compel us to permit the demonstration as proposed, including display of the swastika.

"It is firmly settled that under our Constitution the public expression of ideas may not be prohibited merely because the ideas are themselves offensive to some of their hearers" . . . and it is entirely clear that the wearing of distinctive clothing can be symbolic expression of a thought or philosophy. The symbolic expression of thought falls within the free speech clause of the first amendment . . . and the plaintiff village has the heavy burden of justifying the imposition of a prior restraint upon defendants' right to freedom of speech. . . . The village of Skokie seeks to meet this burden by application of the "fighting words" doctrine first enunciated in *Chaplinsky v. New Hampshire* (1942). . . . That doctrine was designed to permit punishment of extremely hostile personal communication likely to cause immediate physical response, "no words being 'forbidden

373 N.E. 2d 21 (Ill. 1978).

except such as have a direct tendency to cause acts of violence by the persons to whom, individually, the remark is addressed.'" . . . In *Cohen* the Supreme Court restated the description of fighting words as "those personally abusive epithets which, when addressed to the ordinary citizen, are, as a matter of common knowledge, inherently likely to provoke violent reaction." . . . Plaintiff urges, and the appellate court has held, that the exhibition of the Nazi symbol, the swastika, addresses to ordinary citizens a message which is tantamount to fighting words. Plaintiff further asks this court to extend *Chaplinsky,* which upheld a statute punishing the use of such words, and hold that the fighting-words doctrine permits a prior restraint on defendants' symbolic speech. In our judgment we are precluded from doing so.

In *Cohen,* defendant's conviction stemmed from wearing a jacket bearing the words "Fuck the Draft" in a Los Angeles County courthouse corridor. The Supreme Court for reasons we believe applicable here refused to find that the jacket inscription constituted fighting words. That court stated:

> The constitutional right of free expression is powerful medicine in a society as diverse and populous as ours. It is designed and intended to remove government restraints from the arena of public discussion, putting the decision as to what views shall be voiced largely into the hands of each of us, in the hope that use of such freedom will ultimately produce a more capable citizenry and more perfect polity and in the belief that no other approach would comport with the premise of individual dignity and choice upon which our political system rests. . . .
>
> To many, the immediate consequence of this freedom may often appear to be only verbal tumult, discord, and even offensive utterance. These are, however, within established limits, in truth necessary side effects of the broader enduring values which the process of open debate permits us to achieve. That the air may at times seem filled with verbal cacophony is, in this sense not a sign of weakness but of strength. We cannot lose sight of the fact that, in what otherwise might seem a trifling and annoying instance of individual distasteful abuse of a privilege, these fundamental societal values are truly implicated. . . . "[S]o long as the means are peaceful, the communication need not meet standards of acceptability." . . .
>
> Against this perception of the constitutional policies involved, we discern certain more particularized considerations that peculiarly call for reversal of this conviction. First, the principle contended for by the State seems inherently boundless. How is one to distinguish this from any other offensive word [emblem]? Surely the State has no right to cleanse public debate to the point where it is grammatically palatable to the most squeamish among us. Yet no readily ascertainable general principle exists for stopping short of that result were we to affirm the judgment below. For, while the particular four letter word [emblem] being litigated here is perhaps more distasteful than most others of its genre, it is nevertheless often true that one man's vulgarity is another's lyric. Indeed, we think it is largely because governmental officials cannot make principled distinctions in this area that the Constitution leaves matters of taste and style so largely to the individual. . . .
>
> Finally, and in the same vein, we cannot indulge the facile assumption that one can forbid particular words without also running a substantial risk of suppressing ideas in the process. Indeed, governments might soon seize upon the censorship of particular words [emblems] as a convenient guise for banning the expression of unpopular views. We have been able, as noted above, to discern little social benefit that might result from running the risk of opening the door to such grave results. . . .

The display of the swastika, as offensive to the principles of a free nation as the memories it recalls may be, is symbolic political speech intended to convey to the public the beliefs of those who display it. It does not, in our opinion, fall within the definition of "fighting words," and that doctrine cannot be used here to overcome the heavy presumption against the constitutional validity of a prior restraint.

Nor can we find that the swastika, while not representing fighting words, is nevertheless so offensive and peace threatening to the public that its display can be enjoined. We do not doubt that the sight of this symbol is abhorrent to the Jewish citizens of Skokie, and that the survivors of the Nazi persecutions, tormented by their recollections, may have strong feelings regarding its display. Yet it is entirely clear that this factor does not justify enjoining defendants' speech. The *Cohen* court spoke to this subject:

> Finally, in arguments before this Court much has been made of the claim that Cohen's distasteful mode of expression was thrust upon unwilling or unsuspecting viewers, and that the State might therefore legitimately

act as it did in order to protect the sensitive from otherwise unavoidable exposure to appellant's crude form of protest. Of course, the mere presumed presence of unwitting listeners or viewers does not serve automatically to justify curtailing all speech capable of giving offense. . . . While this Court has recognized that government may properly act in many situations to prohibit intrusion into the privacy of the home of unwelcome views and ideas which cannot be totally banned from the public dialogue, . . . we have at the same time consistently stressed that "we are often 'captives' outside the sanctuary of the home and subject to objectionable speech." . . . The ability of government, consonant with the Constitution, to shut off discourse solely to protect others from hearing it is, in other words, dependent upon a showing that substantial privacy interests are being invaded in an essentially intolerable manner. Any broader view of this authority would effectively empower a majority to silence dissidents simply as a matter of personal predilections. . . .

Rockwell v. Morris . . . also involved an American Nazi leader, George Lincoln Rockwell, who challenged a bar to his use of a New York City park to hold a public demonstration where anti-Semitic speeches would be made. Although approximately 2½ million Jewish New Yorkers were hostile to Rockwell's message, the court ordered that a permit to speak be granted, stating:

> A community need not wait to be subverted by street riots and storm troopers; but, also, it cannot, by its policemen or commissioners, suppress a speaker, in prior restraint, on the basis of news reports, hysteria, or inference that what he did yesterday, he will do today. Thus, too, if the speaker incites others to immediate unlawful action he may be punished—in a proper case, stopped when disorder actually impends; but this is not to be confused with unlawful action from others who seek unlawfully to suppress or punish the speaker.
>
> So, the unpopularity of views, their shocking quality, their obnoxiousness, and even their alarming impact is not enough. Otherwise, the preacher of any strange doctrine could be stopped; the antiracist himself could be suppressed, if he undertakes to speak in "restricted" areas; and one who asks that public schools be open indiscriminately to all ethnic groups could be lawfully suppressed, if only he choose to speak where persuasion is needed most. . . .

In summary, as we read the controlling Supreme Court opinions, use of the swastika is a symbolic form of free speech entitled to first amendment protections. Its display on uniforms or banners by those engaged in peaceful demonstrations cannot be totally precluded solely because that display may provoke a violent reaction by those who view it. Particularly is this true where, as here, there has been advance notice by the demonstrators of their plans so that they have become, as the complaint alleges, "common knowledge" and those to whom sight of the swastika banner or uniforms would be offense are forewarned and need not view them. A speaker who gives prior notice of his message has not compelled a confrontation with those who voluntarily listen.

As to those who happen to be in a position to be involuntarily confronted with the swastika, the following observations from *Erznoznik v. City of Jacksonville* . . . are appropriate:

> The plain, if at all times disquieting, truth is that in our pluralistic society, constantly proliferating new and ingenious forms of expression, "we are inescapably captive audiences for many purposes." . . . Much that we encounter offends our esthetic, if not our political and moral, sensibilities. Nevertheless, the Constitution does not permit government to decide which types of otherwise protected speech are sufficiently offensive to require protection for the unwilling listener or viewer. Rather, absent the narrow circumstances described above [home intrusion or captive audience], the burden normally falls upon the viewer to "avoid further bombardment of [his] sensibilities simply by averting [his] eyes."

REVIEW AND DISCUSSION QUESTIONS

1. Describe the factual background and legal issues in this case.
2. The Court relied on an earlier U.S. Supreme Court case, *Cohen v. California,* in deciding this one. What happened in *Cohen?* Why does the Court think it is relevant to *Skokie?*

3. What do you think the intention of the Nazis was in deciding to march in Skokie, if not to win political converts to their cause? Should the Court have considered their purposes?
4. Is the "clear and present danger test" relevant to this case? Explain.
5. How does the Court distinguish this case from *Chaplinsky* and the "fighting words" exception?
6. It is sometimes said that this case was rightly decided because for the Court to rule otherwise would put it on a slippery slope of trying to balance the offensiveness of speech against the First Amendment rights of the speaker, with the result that civil rights marches would have been jeopardized. Do you agree? Explain.

Obscenity

Paris Adult Theatre v. Slaton

Defining obscenity has presented the Supreme Court with serious problems. This case accompanied another, *Miller v. California* (1973), in which the justices offered a new definition of obscenity. Revising the basic approach it had followed since its decision in *Roth v. United States* (1957), the Court said in *Miller* that there are three aspects to a determination of obscenity:

> (a) whether the average person, applying contemporary community standards, would find that the work, taken as a whole, appeals to the prurient interest; (b) whether the work depicts or describes, in a patently offensive way, sexual conduct specifically defined by the applicable state law; and (c) whether the work, taken as a whole, lacks serious literary, artistic, political, or scientific value.

The Court went on to say that "our nation is simply too big and too diverse" for a single standard to apply in all fifty states and therefore held that each jury should evaluate the material based on the standards of the individual community in which the case arose. Armed with its new definition, the Court went on in a companion case to explain the reasoning behind its unwillingness to extend constitutional protection to obscenity. That second case, *Paris Adult Theatre v. Slaton,* arose when police prevented two Atlanta movie theaters from showing sexually explicit movies. No sexually explicit material was displayed outside the theater, and a sign at the entrance said "Adult Theatre—You must be 21 and able to prove it. If viewing the nude body offends you, Please Do Not Enter." The local district attorney tried to get two of the films declared obscene because of their graphic depiction of sexual acts. The lower court agreed that the films were obscene but also held that because there was "requisite notice" to the public of their nature and reasonable protection against admittance of minors, the showing of the films was constitutionally protected. The U.S. Supreme Court, however, disagreed.

Chief Justice Burger, speaking for a majority of the justices, delivered the Supreme Court's verdict that the state is permitted to ban the public exhibition of obscene material to maintain a decent society. In a dissenting opinion, Justice Brennan argues that it is impossible to define "obscenity" satisfactorily.

Mr. Chief Justice Burger: It should be clear from the outset that we do not undertake to tell the States what they must do, but rather to define the area in which they may chart their own course in dealing with obscene material. This Court has consistently held that obscene material is not protected by the First Amendment as a limitation on the state police power by virtue of the Fourteenth Amendment. . . .

---

413 U.S. 49 (1973).

We categorically disapprove the theory, apparently adopted by the trial judge, that obscene, pornographic films acquire constitutional immunity from state regulation simply because they are exhibited for consenting adults only. . . .

In particular, we hold that there are legitimate state interests at stake in stemming the tide of commercialized obscenity, even assuming it is feasible to enforce effective safeguards against exposure to juveniles and to passersby. Rights and interests "other than those of the advocates are involved." These include the interest of the public in the quality of life and the total community environment; the tone of commerce in the great city centers, and, possibly, the public safety itself. The Hill-Lind Minority Report of the Commission on Obscenity and Pornography indicates that there is at least an arguable correlation between obscene material and crime. Quite apart from sex crimes, however, there remains one problem of large proportions aptly described by Professor Bickel: "It concerns the tone of the society, the mode, or to use terms that have perhaps greater currency, the style and quality of life, now and in the future. A man may be entitled to read an obscene book in his room, or expose himself indecently there. . . . We should protect his privacy. But if he demands a right to obtain the books and pictures he wants in the market, and to foregather in public places—discreet, if you will, but accessible to all—with others who share his tastes, *then to grant him his right is to affect the world about the rest of us, and to impinge on other privacies.* Even supposing that each of us can, if he wishes, effectively avert the eye and stop the ear (which, in truth, we cannot), what is commonly read and seen and heard and done intrudes upon us all, want it or not." 22 *The Public Interest* 25–26 (Winter 1971). (Emphasis added.)

As Mr. Chief Justice Warren stated, there is a "right of the Nation and of the States to maintain a decent society. . . ." *Jacobellis v. Ohio.*

But, it is argued, there are no scientific data which conclusively demonstrate that exposure to obscene material adversely affects men and women or their society. It is urged on behalf of the petitioners that, absent such a demonstration, any kind of state regulation is "impermissible." We reject this argument. It is not for us to resolve empirical uncertainties underlying state legislation, save in the exceptional case where that legislation plainly impinges upon rights protected by the Constitution itself. . . .

The sum of experience, including that of the past two decades, affords an ample basis for legislatures to conclude that a sensitive, key relationship of human existence, central to family life, community welfare, and the development of human personality, can be debased and distorted by crass commercial exploitation of sex. Nothing in the Constitution prohibits a State from reaching such a conclusion and acting on it legislatively. . . .

It is also argued that the State has no legitimate interest in "control [of] the moral content of a person's thoughts," and we need not quarrel with this. But we reject the claim that the State of Georgia is here attempting to control the minds or thoughts of those who patronize theaters. Preventing unlimited display or distribution of obscene material, which by definition lacks any serious literary, artistic, political, or scientific value as communication, is distinct from a control of reason and the intellect. . . . Where communication of ideas, protected by the First Amendment, is not involved, or the particular privacy of the home protected by *Stanley* or any of the other "areas or zones" of constitutionally protected privacy, the mere fact that, as a consequence, some human "utterances" or "thoughts" may be incidentally affected does not bar the State from acting to protect legitimate state interests. . . .

The issue in this context goes beyond whether someone, or even the majority, considers the conduct depicted as "wrong" or "sinful." The States have the power to make a morally neutral judgment that public exhibition of obscene material, or commerce in such material, has a tendency to injure the community as a whole, to endanger the public safety, or to jeopardize, in Mr. Chief Justice Warren's words, the States' "right . . . to maintain a decent society." *Jacobellis v. Ohio.* . . .

Mr. Justice Brennan, dissenting: . . . I am convinced that the approach initiated 16 years ago in *Roth v. United States,* 354 U.S. 476 (1957), and culminating in the Court's decision today, cannot bring stability to this area of the law without jeopardizing fundamental First Amendment values, and I have concluded that the time has come to make a significant departure from that approach. . . .

The essence of our problem in the obscenity area is that we have been unable to provide "sensitive tools" to separate obscenity from other sexually oriented but constitutionally protected speech, so that efforts to suppress the former do not spill over into the suppression of the latter. . . .

To be sure, five members of the Court did agree in *Roth* that obscenity could be determined by asking "whether to the average person, applying contemporary community standards, the dominant theme of the material taken as a whole appeals to prurient interest." . . . But agreement on that test—achieved in the abstract and without reference to the particular material before the Court—was, to say the least, short lived. . . .

Today a majority of the Court offers a slightly altered formulation of the basic *Roth* test, while leaving entirely unchanged the underlying approach.

Our experience with the *Roth* approach has certainly taught us that the outright suppression of obscenity cannot be reconciled with the fundamental principles of the First and Fourteenth Amendments. For we have failed to formulate a standard that sharply distinguishes protected from unprotected speech. . . . By disposing of cases through summary reversal or denial of certiorari we have deliberately and effectively obscured the rationale underlying the decisions. It comes as no surprise that judicial attempts to follow our lead conscientiously have often ended in hopeless confusion. . . .

The vagueness of the standards in the obscenity area produces a number of separate problems, and any improvement must rest on an understanding that the problems are to some extent distinct. First, a vague statute fails to provide adequate notice to persons who are engaged in the type of conduct that the statute could be thought to proscribe. The Due Process Clause of the Fourteenth Amendment requires that all criminal laws provide fair notice of "what the State commands or forbids." *Lanzetta v. New Jersey*. . . .

In addition to problems that arise when any criminal statute fails to afford fair notice of what it forbids, a vague statute in the areas of speech and press creates a second level of difficulty. . . .

[R]ecognizing the inherent vagueness of any definition of obscenity, we have held that the definition of obscenity must be drawn as narrowly as possible so as to minimize the interference with protected expression. . . .

The problems of fair notice and chilling protected speech are very grave standing alone. But it does not detract from their importance to recognize that a vague statute in this area creates a third, although admittedly more subtle, set of problems. These problems concern the institutional stress that inevitably results where the line separating protected from unprotected speech is excessively vague. In *Roth* we conceded that "there may be marginal cases in which it is difficult to determine the side of the line on which a particular fact situation falls. . . ." Our subsequent experience demonstrates that almost every case is "marginal." And since the "margin" marks the point of separation between protected and unprotected speech, we are left with a system in which almost every obscenity case presents a constitutional question of exceptional difficulty. . . .

Our experience since *Roth* requires us not only to abandon the effort to pick out obscene materials on a case-by-case basis, but also to reconsider a fundamental postulate of *Roth:* that there exists a definable class of sexually oriented expression that may be totally suppressed by the Federal and State Governments. Assuming that such a class of expression does in fact exist, I am forced to conclude that the concept of "obscenity" cannot be defined with sufficient specificity and clarity to provide fair notice to persons who create and distribute sexually oriented materials, to prevent substantial erosion of protected speech as a byproduct of the attempt to suppress unprotected speech, and to avoid very costly institutional harms.

REVIEW AND DISCUSSION QUESTIONS

1. Describe the three elements of the definition of obscenity outlined in the *Miller* case.
2. Explain the basis of Justice Burger's opinion upholding the Georgia law.
3. Why does Justice Brennan object to Burger's conclusions?
4. How would Mill analyze this case?

Free Speech at School

Morse v. Frederick

In the *Tinker* case in 1969 the Supreme Court held that students do not lose their constitutional rights at school, but it also stated that school officials may restrict speech that substantially disrupts the work and discipline of the school. Since then the Court has sometimes struggled to determine when exactly schools can regulate student speech. In this case, Joseph Frederick was suspended from school for unfurling a banner bearing the words "BONG HiTS 4 JESUS" at a parade that his school had allowed students to leave class to attend. Writing for the Court, Chief Justice Roberts argues that given the serious dangers of illegal drug use, school officials acted permissibly in ordering students to take down a banner that appeared to advocate drug use and in disciplining Frederick for refusing to do so. However, Roberts's opinion provoked two very different concurring opinions (by Justices Thomas and Alito), one opinion concurring in part and dissenting in part (by Justice Breyer), and a dissent opinion (by Justice Stevens).

Chief Justice Roberts delivered the opinion of the Court: At a school-sanctioned and school-supervised event, a high school principal saw some of her students unfurl a large banner conveying a message she reasonably regarded as promoting illegal drug use. Consistent with established school policy prohibiting such messages at school events, the principal directed the students to take down the banner. One student—among those who had brought the banner to the event—refused to do so. The principal confiscated the banner and later suspended the student. The Ninth Circuit held that the principal's actions violated the First Amendment, and that the student could sue the principal for damages.

Our cases make clear that students do not "shed their constitutional rights to freedom of speech or expression at the schoolhouse gate." *Tinker* v. *Des Moines Independent Community School Dist.* (1969). At the same time, we have held that "the constitutional rights of students in public school are not automatically coextensive with the rights of adults in other settings" and that the rights of students "must be 'applied in light of the special characteristics of the school environment.'" Consistent with these principles, we hold that schools may take steps to safeguard those entrusted to their care from speech that can reasonably be regarded as encouraging illegal drug use. We conclude that the school officials in this case did not violate the First Amendment by confiscating the pro-drug banner and suspending the student responsible for it.

I

On January 24, 2002, the Olympic Torch Relay passed through Juneau, Alaska, on its way to the winter games in Salt Lake City, Utah. The torchbearers were to proceed along a street in front of Juneau-Douglas High School (JDHS) while school was in session. Petitioner Deborah Morse, the school principal, decided to permit staff and students to participate in the Torch Relay as an approved social event or class trip. Students were allowed to leave class to observe the relay from either side of the street. Teachers and administrative officials monitored the students' actions.

Respondent Joseph Frederick, a JDHS senior, was late to school that day. When he arrived, he joined his friends . . . across the street from the school to watch the event. . . . As the torchbearers and camera crews passed by, Frederick and his friends unfurled a 14-foot banner bearing the phrase: "BONG HiTS 4 JESUS." The large banner was easily readable by the students on the other side of the street.

Principal Morse immediately crossed the street and demanded that the banner be taken down.

127 S.Ct. 2618 (2007).

Everyone but Frederick complied. Morse confiscated the banner and told Frederick to report to her office, where she suspended him for 10 days. Morse later explained that she told Frederick to take the banner down because she thought it encouraged illegal drug use, in violation of established school policy. . . .

Frederick administratively appealed his suspension, but the Juneau School District Superintendent upheld it, limiting it to time served (8 days). . . . Frederick then filed suit under 42 U. S. C. §1983, alleging that the school board and Morse had violated his First Amendment rights. . . . The District Court granted summary judgment for the school board and Morse, ruling that they were entitled to qualified immunity and that they had not infringed Frederick's First Amendment rights. The court found that Morse reasonably interpreted the banner as promoting illegal drug use—a message that "directly contravened the Board's policies relating to drug abuse prevention." . . .

The Ninth Circuit reversed. Deciding that Frederick acted during a "school-authorized activ-it[y]," and "proceed[ing] on the basis that the banner expressed a positive sentiment about marijuana use," the court nonetheless found a violation of Frederick's First Amendment rights because the school punished Frederick without demonstrating that his speech gave rise to a "risk of substantial disruption." The court further concluded that Frederick's right to display his banner was so "clearly established" that a reasonable principal in Morse's position would have understood that her actions were unconstitutional, and that Morse was therefore not entitled to qualified immunity.

We granted certiorari on two questions: whether Frederick had a First Amendment right to wield his banner, and, if so, whether that right was so clearly established that the principal may be held liable for damages. We resolve the first question against Frederick, and therefore have no occasion to reach the second.

II

At the outset, we reject Frederick's argument that this is not a school speech case—as has every other authority to address the question. The event occurred during normal school hours. It was sanctioned by Principal Morse "as an approved social event or class trip," and the school district's rules expressly provide that pupils in "approved social events and class trips are subject to district rules for student conduct." . . .

III

The message on Frederick's banner is cryptic. It is no doubt offensive to some, perhaps amusing to others. To still others, it probably means nothing at all. Frederick himself claimed "that the words were just nonsense meant to attract television cameras." But Principal Morse thought the banner would be interpreted by those viewing it as promoting illegal drug use, and that interpretation is plainly a reasonable one. . . .

The pro-drug interpretation of the banner gains further plausibility given the paucity of alternative meanings the banner might bear. The best Frederick can come up with is that the banner is "meaningless and funny." Gibberish is surely a possible interpretation of the words on the banner, but it is not the only one, and dismissing the banner as meaningless ignores its undeniable reference to illegal drugs. . . .

IV

The question thus becomes whether a principal may, consistent with the First Amendment, restrict student speech at a school event, when that speech is reasonably viewed as promoting illegal drug use. We hold that she may.

In *Tinker*, this Court made clear that "First Amendment rights, applied in light of the special characteristics of the school environment, are available to teachers and students." *Tinker* involved a group of high school students who decided to wear black armbands to protest the Vietnam War. School officials learned of the plan and then adopted a policy prohibiting students from wearing armbands. When several students nonetheless wore armbands to school, they were suspended. The students sued, claiming that their First Amendment rights had been violated, and this Court agreed.

Tinker held that student expression may not be suppressed unless school officials reasonably conclude that it will "materially and substantially disrupt the work and discipline of the school." . . .

This Court's next student speech case was *Fraser*. Matthew Fraser was suspended for delivering a speech before a high school assembly in which he employed what this Court called "an elaborate, graphic, and explicit sexual metaphor." Analyzing the case under *Tinker*, the District Court and Court of Appeals found no disruption, and therefore no basis for disciplining Fraser. This Court reversed, holding that the "School District acted entirely within its permissible authority in imposing sanctions upon Fraser in response to his offensively lewd and indecent speech."

. . . For present purposes, it is enough to distill from *Fraser* two basic principles. First, *Fraser*'s holding demonstrates that "the constitutional rights of students in public school are not automatically coextensive with the rights of adults in other settings." Had Fraser delivered the same speech in a public forum outside the school context, it would have been protected. In school, however, Fraser's First Amendment rights were circumscribed "in light of the special characteristics of the school environment." Second, *Fraser* established that the mode of analysis set forth in *Tinker* is not absolute. Whatever approach *Fraser* employed, it certainly did not conduct the "substantial disruption" analysis prescribed by *Tinker*. . . .

The "special characteristics of the school environment," *Tinker*, 393 *U. S.*, at 506, and the governmental interest in stopping student drug abuse—reflected in the policies of Congress and myriad school boards, including JDHS—allow schools to restrict student expression that they reasonably regard as promoting illegal drug use. *Tinker* warned that schools may not prohibit student speech because of "undifferentiated fear or apprehension of disturbance" or "a mere desire to avoid the discomfort and unpleasantness that always accompany an unpopular viewpoint." The danger here is far more serious and palpable. The particular concern to prevent student drug abuse at issue here, embodied in established school policy, extends well beyond an abstract desire to avoid controversy. . . .

School principals have a difficult job, and a vitally important one. When Frederick suddenly and unexpectedly unfurled his banner, Morse had to decide to act—or not act—on the spot. It was reasonable for her to conclude that the banner promoted illegal drug use—in violation of established school policy—and that failing to act would send a powerful message to the students in her charge, including Frederick, about how serious the school was about the dangers of illegal drug use. The First Amendment does not require schools to tolerate at school events student expression that contributes to those dangers.

The judgment of the United States Court of Appeals for the Ninth Circuit is reversed, and the case is remanded for further proceedings consistent with this opinion.

It is so ordered.

JUSTICE THOMAS, concurring: The Court today decides that a public school may prohibit speech advocating illegal drug use. I agree and therefore join its opinion in full. I write separately to state my view that the standard set forth in *Tinker* v. *Des Moines Independent Community School Dist.*, 393 U. S. 503 (1969), is without basis in the Constitution.

I

. . . In my view, the history of public education suggests that the First Amendment, as originally understood, does not protect student speech in public schools. Although colonial schools were exclusively private, public education proliferated in the early 1800's. By the time the States ratified the Fourteenth Amendment, public schools had become relatively common.

If students in public schools were originally understood as having free-speech rights, one would have expected 19th-century public schools to have respected those rights and courts to have enforced them. They did not. . . .

II

Tinker effected a sea change in students' speech rights, extending them well beyond traditional bounds. . . . Of course, *Tinker*'s reasoning conflicted

with the traditional understanding of the judiciary's role in relation to public schooling, a role limited by *in loco parentis.* Perhaps for that reason, the Court has since scaled back *Tinker*'s standard, or rather set the standard aside on an ad hoc basis. . . .

Today, the Court creates another exception. In doing so, we continue to distance ourselves from *Tinker*, but we neither overrule it nor offer an explanation of when it operates and when it does not. . . .

III

In light of the history of American public education, it cannot seriously be suggested that the First Amendment "freedom of speech" encompasses a student's right to speak in public schools. Early public schools gave total control to teachers, who expected obedience and respect from students. And courts routinely deferred to schools' authority to make rules and to discipline students for violating those rules. Several points are clear: (1) under *in loco parentis*, speech rules and other school rules were treated identically; (2) the *in loco parentis* doctrine imposed almost no limits on the types of rules that a school could set while students were in school; and (3) schools and teachers had tremendous discretion in imposing punishments for violations of those rules. . . .

I join the Court's opinion because it erodes *Tinker*'s hold in the realm of student speech, even though it does so by adding to the patchwork of exceptions to the *Tinker* standard. I think the better approach is to dispense with *Tinker* altogether, and given the opportunity, I would do so.

Justice Alito, with whom Justice Kennedy joins, concurring: I join the opinion of the Court on the understanding that (a) it goes no further than to hold that a public school may restrict speech that a reasonable observer would interpret as advocating illegal drug use and (b) it provides no support for any restriction of speech that can plausibly be interpreted as commenting on any political or social issue, including speech on issues such as "the wisdom of the war on drugs or of legalizing marijuana for medicinal use." . . .

The public schools are invaluable and beneficent institutions, but they are, after all, organs of the State. When public school authorities regulate student speech, they act as agents of the State; they do not stand in the shoes of the students' parents. . . . It is therefore wrong to treat public school officials, for purposes relevant to the First Amendment, as if they were private, nongovernmental actors standing *in loco parentis.*

For these reasons, any argument for altering the usual free speech rules in the public schools . . . must . . . be based on some special characteristic of the school setting. The special characteristic that is relevant in this case is the threat to the physical safety of students. School attendance can expose students to threats to their physical safety that they would not otherwise face. . . .

In most settings, the First Amendment strongly limits the government's ability to suppress speech on the ground that it presents a threat of violence. But due to the special features of the school environment, school officials must have greater authority to intervene before speech leads to violence. And, in most cases, *Tinker's* "substantial disruption" standard permits school officials to step in before actual violence erupts.

Speech advocating illegal drug use poses a threat to student safety that is just as serious, if not always as immediately obvious. As we have recognized in the past and as the opinion of the Court today details, illegal drug use presents a grave and in many ways unique threat to the physical safety of students. I therefore conclude that the public schools may ban speech advocating illegal drug use. But I regard such regulation as standing at the far reaches of what the First Amendment permits. I join the opinion of the Court with the understanding that the opinion does not endorse any further extension.

Justice Breyer, concurring in the judgment in part and dissenting in part: This Court need not and should not decide this difficult First Amendment issue on the merits. Rather, I believe that it should simply hold that qualified immunity bars the student's claim for monetary damages and say no more.

I

Resolving the First Amendment question presented in this case is, in my view, unwise and unnecessary. In part that is because the question focuses upon specific content narrowly defined: May a school board punish students for speech that advocates drug use and, if so, when? At the same time, the underlying facts suggest that Principal Morse acted as she did not simply because of the specific content and viewpoint of Joseph Frederick's speech but also because of the surrounding context and manner in which Frederick expressed his views. To say that school officials might reasonably prohibit students during school-related events from unfurling 14-foot banners (with any kind of irrelevant or inappropriate message) designed to attract attention from television cameras seems unlikely to undermine basic First Amendment principles. But to hold, as the Court does, that "schools may take steps to safeguard those entrusted to their care from speech that can reasonably be regarded as encouraging illegal drug use" (and that "schools" may "restrict student expression that they reasonably regard as promoting illegal drug use") is quite a different matter. This holding, based as it is on viewpoint restrictions, raises a host of serious concerns.

One concern is that, while the holding is theoretically limited to speech promoting the use of illegal drugs, it could in fact authorize further viewpoint-based restrictions. Illegal drugs, after all, are not the only illegal substances. What about encouraging the underage consumption of alcohol? Moreover, it is unclear how far the Court's rule regarding drug advocacy extends. What about a conversation during the lunch period where one student suggests that glaucoma sufferers should smoke marijuana to relieve the pain? What about deprecating commentary about an antidrug film shown in school? And what about drug messages mixed with other, more expressly political, content? If, for example, Frederick's banner had read "LEGALIZE BONG HiTS," he might be thought to receive protection from the majority's rule, which goes to speech "encouraging illegal drug use." But speech advocating change in drug laws might also be perceived of as promoting the disregard of existing drug laws. . . .

All of this is to say that a decision on the underlying First Amendment issue . . . is both difficult and unusually portentous. And that is a reason for us *not to decide* the issue unless we must.

In some instances, it is appropriate to decide a constitutional issue in order to provide "guidance" for the future. But I cannot find much guidance in today's decision. The Court makes clear that school officials may "restrict" student speech that promotes "illegal drug use" and that they may "take steps" to "safeguard" students from speech that encourages "illegal drug use." Beyond "steps" that prohibit the unfurling of banners at school outings, the Court does not explain just what those "restrict[ions]" or those "steps" might be.

Nor, if we are to avoid the risk of interpretations that are too broad or too narrow, is it easy to offer practically valuable guidance. Students will test the limits of acceptable behavior in myriad ways better known to school teachers than to judges; school officials need a degree of flexible authority to respond to disciplinary challenges; and the law has always considered the relationship between teachers and students special. Under these circumstances, the more detailed the Court's supervision becomes, the likely its law will engender further disputes among teachers and students. Consequently, larger numbers of those disputes will likely make their way from the schoolhouse to the courthouse. Yet no one wishes to substitute courts for school boards, or to turn the judge's chambers into the principal's office.

In order to avoid resolving the fractious underlying constitutional question, we need only decide a different question that this case presents, the question of "qualified immunity." The principle of qualified immunity fits this case perfectly and, by saying so, we would diminish the risk of bringing about the adverse consequences I have identified. More importantly, we should also adhere to a basic constitutional obligation by avoiding unnecessary decision of constitutional questions.

II

The defense of "qualified immunity" requires courts to enter judgment in favor of a government employee unless the employee's conduct violates "clearly established statutory or constitutional

rights of which a reasonable person would have known." The defense is designed to protect "all but the plainly incompetent or those who knowingly violated the law." Qualified immunity applies here and entitles Principal Morse to judgment on [her behalf] . . . because she did not clearly violate the law during her confrontation with the student.

Justice Stevens, with whom Justice Souter and Justice Ginsburg join, dissenting: . . . On January 24, 2002, the Olympic Torch Relay gave those Alaska residents a rare chance to appear on national television. As Joseph Frederick repeatedly explained, he did not address the curious message—"BONG HiTS 4 JESUS"—to his fellow students. He just wanted to get the camera crews' attention. Moreover, concern about a nationwide evaluation of the conduct of the JDHS student body would have justified the principal's decision to remove an attention-grabbing 14-foot banner, even if it had merely proclaimed "Glaciers Melt!"

I agree with the Court that the principal should not be held liable for pulling down Frederick's banner. I would hold, however, that the school's interest in protecting its students from exposure to speech "reasonably regarded as promoting illegal drug use" cannot justify disciplining Frederick for his attempt to make an ambiguous statement to a television audience simply because it contained an oblique reference to drugs. The First Amendment demands more, indeed, much more.

The Court holds otherwise only after laboring to establish two uncontroversial propositions: first, that the constitutional rights of students in school settings are not coextensive with the rights of adults, and second, that deterring drug use by schoolchildren is a valid and terribly important interest. . . . But it is a gross non sequitur to draw from these two unremarkable propositions the remarkable conclusion that the school may suppress student speech that was never meant to persuade anyone to do anything.

In my judgment, the First Amendment protects student speech if the message itself neither violates a permissible rule nor expressly advocates conduct that is illegal and harmful to students. This nonsense banner does neither, and the Court does serious violence to the First Amendment in upholding—indeed, lauding—a school's decision to punish Frederick for expressing a view with which it disagreed.

I

. . . Two cardinal First Amendment principles animate . . . the Court's opinion in *Tinker*. . . . First, censorship based on the content of speech, particularly censorship that depends on the viewpoint of the speaker, is subject to the most rigorous burden of justification. . . . Second, punishing someone for advocating illegal conduct is constitutional only when the advocacy is likely to provoke the harm that the government seeks to avoid.

However necessary it may be to modify those principles in the school setting, *Tinker* affirmed their continuing vitality. . . .

Yet today the Court fashions a test that trivializes the two cardinal principles upon which *Tinker* rests. ("[S]chools [may] restrict student expression that they reasonably regard as promoting illegal drug use.") The Court's test invites stark viewpoint discrimination. In this case, for example, the principal has unabashedly acknowledged that she disciplined Frederick because she disagreed with the pro-drug viewpoint she ascribed to the message on the banner—a viewpoint, incidentally, that Frederick has disavowed, see *id.*, at 28. Unlike our recent decision in *Tennessee Secondary School Athletic Assn.* v. *Brentwood Academy* (2007), the Court's holding in this case strikes at "the heart of the First Amendment" because it upholds a punishment meted out on the basis of a listener's disagreement with her understanding (or, more likely, misunderstanding) of the speaker's viewpoint. "If there is a bedrock principle underlying the First Amendment, it is that the Government may not prohibit the expression of an idea simply because society finds the idea itself offensive or disagreeable." *Texas* v. *Johnson* (1989).

It is also perfectly clear that "promoting illegal drug use" comes nowhere close to proscribable "incitement to imminent lawless action." Encouraging drug use might well increase the

likelihood that a listener will try an illegal drug, but that hardly justifies censorship. . . . No one seriously maintains that drug advocacy (much less Frederick's ridiculous sign) comes within the vanishingly small category of speech that can be prohibited because of its feared consequences. . . .

II

The Court rejects outright these twin foundations of *Tinker* because, in its view, the unusual importance of protecting children from the scourge of drugs supports a ban on all speech in the school environment that promotes drug use. Whether or not such a rule is sensible as a matter of policy, carving out pro-drug speech for uniquely harsh treatment finds no support in our case law and is inimical to the values protected by the First Amendment. . . .

III

Although this case began with a silly, nonsensical banner, it ends with the Court inventing out of whole cloth a special First Amendment rule permitting the censorship of any student speech that mentions drugs, at least so long as someone could perceive that speech to contain a latent pro-drug message. . . .

Even in high school, a rule that permits only one point of view to be expressed is less likely to produce correct answers than the open discussion of countervailing views. In the national debate about a serious issue, it is the expression of the minority's viewpoint that most demands the protection of the First Amendment. Whatever the better policy may be, a full and frank discussion of the costs and benefits of the attempt to prohibit the use of marijuana is far wiser than suppression of speech because it is unpopular.

I respectfully dissent.

REVIEW AND DISCUSSION QUESTIONS

1. Explain the facts and legal background of the case.
2. What was decided in the earlier cases of *Tinker* and *Fraser*, and how do those cases bear on the present case? In light of those precedents, explain why you agree or disagree that the First Amendment permits schools to censor speech that advocates the use of illegal drugs.
3. Did Frederick's banner, in fact, advocate drug use?
4. Why does Justice Thomas believe that *Tinker* should be overturned? How persuasive do you find his reasoning?
5. Why did Justice Alito write a separate, concurring opinion? Would he agree or disagree with Justice Thomas?
6. Why does Justice Breyer believe that the Court should avoid, if it can, deciding the underlying First Amendment issue? Do you agree or disagree with his reasoning?
7. Justice Breyer argues that the Court could decide the issue in favor of the principal on the basis of "qualified immunity." Explain.
8. On what grounds does Justice Stevens dissent? Is he right that this is a case of censorship based on content?
9. How would you have decided this case?

CHAPTER 22

Equality

The U.S. Constitution's Fourteenth Amendment demands that government respect the equality of persons, stating that everyone must be provided "equal protection of the laws." It also might be argued that equality is implicit elsewhere in the Constitution, for instance in the requirements that government not infringe fundamental liberties, including the freedoms of religious practice and speech and the right to due process under the law. But what, precisely, is required of government if it is to respect equality? The readings and cases in this section weigh a range of topics, including racial segregation, the internment of Japanese Americans during World War II, affirmative action, and gender discrimination.

Separate but Equal

Plessy v. Ferguson

Three constitutional amendments were passed after the Civil War. The Thirteenth Amendment abolished slavery; the Fifteenth Amendment promised that the right to vote will not be denied on account of race, color or previous condition of servitude; and the Fourteenth Amendment stated that the states cannot deny "equal protection" of the laws, abridge the "privileges or immunities of citizens of the United States" nor deprive citizens of life, liberty or property without "due process of law." Some early cases indicated that the Court might work to enforce these ideals, but the Court backtracked, and soon throughout the South radicals committed to removing federal troops, maintaining racial segregation, and disenfranchising blacks were in control. In this famous 1896 case, a black man had been arrested and prosecuted for entering a railroad coach reserved for whites and refusing to move to one reserved for blacks. The Supreme Court upheld his conviction under Louisiana law, endorsing the general principle of "separate but equal," namely, that segregation of the races is constitutionally permissible, provided that the facilities offered blacks are substantially equal to those given to whites.

163 U.S. 537 (1896).

Mr. Justice Brown: The distinction between laws interfering with the political equality of the negro and those requiring the separation of the two races in schools, theatres and railway carriages has been frequently drawn by this court. . . . [W]e think the enforced separation of the races, as applied to the internal commerce of the State, neither abridges the privileges or immunities of the colored man, deprives him of his property without due process of law, nor denies him the equal protection of the laws, within the meaning of the Fourteenth Amendment. . . . [I]t is suggested . . . that the same argument that will justify the state legislature in requiring railways to provide separate accommodations for the two races will also authorize them to require separate cars to be provided for people whose hair is of a certain color, or who are aliens, or who belong to certain nationalities, or to enact laws requiring colored people to walk upon one side of the street, and white people upon the other, or requiring white men's houses to be painted white, and colored men's black, or their vehicles or business signs to be of different colors, upon the theory that one side of the street is as good as the other, or that a house or vehicle of one color is as good as one of another color. The reply to all this is that every exercise of the police power must be reasonable, and extend only to such laws as are enacted in good faith for the promotion for the public good, and not for the annoyance or oppression of a particular class. . . .

So far, then, as a conflict with the Fourteenth Amendment is concerned, the case reduces itself to the question whether the statute of Louisiana is a reasonable regulation, and with respect to this there must necessarily be a large discretion on the part of the legislature. In determining the question of reasonableness it is at liberty to act with reference to the established usages, customs and traditions of the people, and with a view to the promotion of their comfort, and the preservation of the public peace and good order. Gauged by this standard, we cannot say that a law which authorizes or even requires the separation of the two races in public conveyances is unreasonable, or more obnoxious to the Fourteenth Amendment than the acts of Congress requiring separate schools for colored children in the District of Columbia, the constitutionality of which does not seem to have been questioned, or the corresponding acts of state legislatures.

We consider the underlying fallacy of the plaintiff's argument to consist in the assumption that the enforced separation of the two races stamps the colored race with a badge of inferiority. If this be so, it is not by reason of anything found in the act, but solely because the colored race chooses to put that construction upon it. The argument necessarily assumes that if, as has been more than once the case, and is not unlikely to be so again, the colored race should become the dominant power in the state legislature, and should enact a law in precisely similar terms, it would thereby relegate the white race to an inferior position. We imagine that the white race, at least, would not acquiesce in this assumption. The argument also assumes that social prejudices may be overcome by legislation, and that equal rights cannot be secured to the negro except by an enforced commingling of the two races. We cannot accept this proposition. If the two races are to meet upon terms of social equality, it must be the result of natural affinities, a mutual appreciation of each other's merits and a voluntary consent of individuals. . . . Legislation is powerless to eradicate racial instincts or to abolish distinctions based upon physical differences, and the attempt to do so can only result in accentuating the difficulties of the present situation. If the civil and political rights of both races be equal one cannot be inferior to the other civilly or politically. If one race be inferior to the other socially, the Constitution of the United States cannot put them upon the same plane.

Mr. Justice Harlan, dissenting: The white race deems itself to be the dominant race in this country. And so it is, in prestige, in achievements, in education, in wealth and in power. So, I doubt not, it will continue to be for all time, if it remains true to its great heritage and holds fast to the principles of constitutional liberty. But in view of the Constitution, in the eye of the law, there is in this country no superior, dominant, ruling class of citizens. There is no caste here. Our Constitution is color-blind, and neither knows nor tolerates classes among citizens. In respect of civil rights, all citizens are equal before the law. The humblest is the peer

of the most powerful. The law regards man as man, and takes no account of his surroundings or of his color when his civil rights as guaranteed by the supreme law of the land are involved. It is, therefore, to be regretted that this high tribunal, the final expositor of the fundamental law of the land, has reached the conclusion that it is competent for a State to regulate the enjoyment by citizens of their civil rights solely upon the basis of race.

In my opinion, the judgment this day rendered will, in time, prove to be quite as pernicious as the decision made by this tribunal in the *Dred Scott* case [1857]. . . . The destinies of the two races, in this country, are indissolubly linked together, and the interests of both require that the common government of all shall not permit the seeds of race hate to be planted under the sanction of law. What can more certainly arouse race hate, what more certainly create and perpetuate a feeling of distrust between these races, than state enactments, which, in fact, proceed on the ground that colored citizens are so inferior and degraded that they cannot be allowed to sit in public coaches occupied by white citizens? That, as all will admit, is the real meaning of such legislation as was enacted in Louisiana.

The sure guarantee of the peace and security of each race is the clear, distinct, unconditional recognition by our governments, National and State, of every right that inheres in civil freedom, and of the equality before the law of all citizens of the United States without regard to race. State enactments regulating the enjoyment of civil rights, upon the basis of race and cunningly devised to defeat legitimate results of the war, under the pretence of recognizing equality of rights, can have no other result than to render permanent peace impossible, and to keep alive a conflict of races, the continuance of which must do harm to all concerned. This question is not met by the suggestion that social equality cannot exist between the white and black races in this country . . . for social equality no more exists between two races when travelling in a passenger coach or a public highway than when members of the same races sit by each other in a street car or in the jury box, or stand or sit with each other in a political assembly, or when they use in common the streets of a city or town. . . .

The arbitrary separation of citizens, on the basis of race, while they are on a public highway, is a badge of servitude wholly inconsistent with the civil freedom and the equality before the law established by the Constitution. It cannot be justified upon any legal grounds. . . .

We boast of the freedom enjoyed by our people above all other peoples. But it is difficult to reconcile that boast with a state of the law which, practically, puts the brand of servitude and degradation upon a large class of our fellow-citizens, our equals before the law. The thin disguise of "equal" accommodations for passengers in railroad coaches will not mislead any one, not atone for the wrong this day done.

REVIEW AND DISCUSSION QUESTIONS

1. Describe the basic argument offered by Justice Brown in defense of the Louisiana law's constitutionality.
2. What does he see as the "underlying fallacy" of Plessy's argument?
3. Has history proved Justice Brown right in his claim about the limits of legislation?
4. Justice Harlan wrote a famous dissent. Explain briefly his reasoning.

The Great School Desegregation Case

Brown v. Board of Education

In a few cases during the 1930s, the Supreme Court began to show some concern for the way states habitually treated black citizens. In one case, *Brown v. Mississippi* (1936), three blacks had been arrested for murder. The deputy sheriff hanged one of them by the neck from a tree, occasionally lowering him only to ask if he were ready to confess, which he eventually did. All three were found guilty after a brief trial before an all-white jury, without the benefit of a lawyer. Even the rope scars did not sway the jurors, who sentenced all three to death. The Supreme Court reversed the conviction, saying that the defendants had been denied their rights under the due process clause of the Fourteenth Amendment.

But by far the biggest victories came in the area of education. From its founding in 1909 the National Association for the Advancement of Colored People (NAACP) fought to overturn *Plessy v. Ferguson.* Then in 1938 a young civil rights lawyer named Thurgood Marshall was named director of the NAACP Legal Defense and Education Fund, and for the next eighteen years he and others won an extraordinary series of legal victories that slowly but surely undermined the legal basis of state-enforced segregation. In 1954 the NAACP attacked *Plessy* directly in a case that involved the segregated educational systems of four states plus the District of Columbia. After hearing the initial arguments, the Justices tentatively voted five to four in favor of the NAACP's position. Chief Justice Vinson soon died, and President Eisenhower appointed Earl Warren to replace him. Warren worked hard to persuade the remaining dissenters, winning over the last holdout, Justice Reed, by arguing that unanimity was crucial if the order was to be carried out and that his love of the Court should take precedence over his personal views.

The opinion was widely hated; calls for Warren's impeachment could be seen on billboards throughout the South. Prince Edward County, Virginia went so far as to close its public schools rather than integrate them, and it was not until the Civil Rights Act of 1964 along with more aggressive efforts by the Courts to enforce its decision in *Brown* that progress was made. Nonetheless, most agree that *Brown v. Board of Education* is among the Supreme Court's greatest achievements.

Chief Justice Warren delivered the opinion of the Court: The plaintiffs contend that segregated public schools are not "equal" and cannot be made "equal," and that hence they are deprived of the equal protection of the laws. . . .

Reargument was largely devoted to the circumstances surrounding the adoption of the Fourteenth Amendment in 1868. . . . This discussion and our own investigation convince us that, although these sources cast some light, it is not enough to resolve the problem with which we are faced. At best, they are inconclusive. The most avid proponents of the post-War Amendments undoubtedly intended them to remove all legal distinctions among "all persons born or naturalized in the United States." Their opponents, just as certainly, were antagonistic to both the letter and the spirit of the Amendments and wished them to have the most limited effect. . . .

An additional reason for the inclusive nature of the Amendment's history, with respect to segregated schools, is the status of public education at that time. In the South, the movement toward free common schools, supported by general taxation, had not yet taken hold. Education of white children was largely in the hands of private groups. Education of Negroes was almost nonexistent, and practically all of the race were illiterate. In fact, any education of Negroes was forbidden by law in some states. . . .

In the first cases in this Court construing the Fourteenth Amendment, decided shortly after its adoption, the Court interpreted it as proscribing all state-imposed discriminations against the Negro

347 U.S. 483 (1954).

race. The doctrine of "separate but equal" did not make its appearance in this Court until 1896 in the case of *Plessy v. Ferguson*, involving not education but transportation. American courts have since labored with the doctrine for over half a century. . . .

In more recent cases, all on the graduate school level, inequality was found in that specific benefits enjoyed by white students were denied to Negro students of the same educational qualifications. . . .

In the instant cases, that question is directly presented. Here, unlike *Sweatt v. Painter* (1950), there are findings below that the Negro and white schools involved have been equalized, or are being equalized, with respect to buildings, curricula, qualifications and salaries of teachers, and other "tangible" factors. Our decision, therefore, cannot turn on merely a comparison of these tangible factors in the Negro and white schools involved in each of the cases. We must look instead to the effect of segregation itself on public education.

In approaching this problem, we cannot turn the clock back to 1868 when the Amendment was adopted, or even to 1896 when *Plessy* was written. We must consider public education in the light of its full development and its present place in American life throughout the Nation. Only in this way can it be determined if segregation in public schools deprives these plaintiffs of the equal protection of the laws.

Today, education is perhaps the most important function of state and local governments. Compulsory school attendance laws and the great expenditures for education both demonstrate our recognition of the importance of education to our democratic society. . . . In these days, it is doubtful that any child may reasonably be expected to succeed in life if he is denied the opportunity of an education. Such an opportunity, where the state has undertaken to provide it, is a right which must be made available to all on equal terms.

We come then to the question presented: Does segregation of children in public schools solely on the basis of race, even though the physical facilities and other "tangible" factors may be equal, deprive the children of the minority group of equal educational opportunities? We believe that it does.

In *Sweatt*, in finding that a segregated law school for Negroes could not provide them equal educational opportunities, this Court relied in large part on "those qualities which are incapable of objective measurement but which make for greatness in a law school." In *McLaurin v. Oklahoma* (1950), the Court, in requiring that a Negro admitted to a white graduate school be treated like all other students, again resorted to intangible considerations: ". . . his ability to study, to engage in discussions and exchange views with other students, and, in general, to learn his profession." Such considerations apply with added force to children in grade and high schools. To separate them from others of similar age and qualifications solely because of their race generates a feeling of inferiority as to their status in the community that may affect their hearts and minds in a way unlikely ever to be undone. The effect of this separation on their educational opportunities was well stated by a finding in the Kansas case by a court which nevertheless felt compelled to rule against the Negro plaintiffs: "Segregation of white and colored children in public schools has a detrimental effect upon the colored children. The impact is greater when it has the sanction of the law; for the policy of separating the races is usually interpreted as denoting the inferiority of the negro group. A sense of inferiority affects the motivation of a child to learn. Segregation with the sanction of law, therefore, has a tendency to [retard] the educational and mental development of Negro children and to deprive them of some of the benefits they would receive in a racial[ly] integrated school system."

Whatever may have been the extent of psychological knowledge at the time of *Plessy*, this finding is amply supported by modern authority. [Here the Court cites various studies, including ones showing that both black and white children prefer white dolls.—EDS.] Any language in *Plessy* contrary to this finding is rejected.

We conclude that in the field of public education the doctrine of "separate but equal" has no place. Separate educational facilities are inherently unequal. Therefore, we hold that the plaintiffs and others similarly situated for whom the actions have been brought are, by reason of the segregation complained of, deprived of the equal protection of the laws guaranteed by the Fourteenth Amendment.

Because these are class actions, because of the wide applicability of this decision, and because of

the great variety of local conditions, the formulation of decrees in these cases presents problems of considerable complexity. On reargument, the consideration of appropriate relief was necessarily subordinated to the primary question—the constitutionality of segregation in public education. We have now announced that such segregation is a denial of the equal protection of the laws.

REVIEW AND DISCUSSION QUESTIONS

1. The Court did not explicitly overrule *Plessy* in *Brown*. Explain just what the Court did do.
2. What is the sociological basis of the Court's decision?
3. Why do you think the Court thought desegregation could be expected to improve the educational opportunities for black children?
4. Judging from the Court's opinion in *Brown* alone, what would you have predicted it would do with cases involving segregated public swimming pools and legal bans on interracial marriage? (Later decisions, in fact, struck these down as well, using *Brown* as precedent.)
5. This has been described as a conservative opinion because the country had no real choice at the time but to desegregate. Describe the social, economic, and international factors that might be used to make that argument.
6. Compare this decision with Justice Harlan's dissent in *Plessy*.
7. Responding to the refusal of school districts to abide by the decision in *Brown*, later courts ordered school busing both within and across district lines to eliminate the vestiges of past discrimination. Are those remedies consistent with this decision? Are they required by *Brown?*
8. There are still many public, "historically black" universities in the United States, most of which were created by states during the time of segregation. Though they allow whites to attend, few whites choose to do so, and the schools remain overwhelmingly black. Should courts now force them to integrate, even though many black colleges resist? Is the fact that blacks choose to go relevant? ("Freedom of choice" plans were struck down by the Court for primary and secondary schools, which described them as a means to avoid integration.)
9. In *Mississippi University for Women v. Hogan* (1982) Justice Sandra Day O'Connor ruled that state-supported women's colleges are inconsistent with the Equal Protection Clause. How might you argue she was right in that case? How would you argue she was wrong?

Wartime Internment of Japanese Americans

Korematsu v. United States

During World War II fear of invasion and of the power of the Japanese army were widespread after Japan conquered much of the Western Pacific, having already destroyed most of the United States naval fleet at Pearl Harbor. At the same time, there was growing pressure along the West Coast of the United States for the military to do something about the minority population of Japanese Americans. Some argued they could aid an invading army; others feared for the welfare of people of Japanese descent at the hands of angry mobs; and, finally, there was also a fair amount of economic advantage to be gained if these successful, hard-working people were forced to sell or otherwise lose their businesses and homes. Although people in both military intelligence and in the Federal Bureau of Investigation disputed the claim that Japanese Americans posed a danger, Franklin Roosevelt nevertheless signed a bill into law

323 U.S. 214 (1944).

allowing the military to establish "military zones" from which any persons could be excluded and that validated the military's already established curfew on both Japanese citizens and American citizens of Japanese descent living on the West Coast.

In *Hirabayashi v. United States* (1943), the Supreme Court had upheld the conviction of a University of Washington student, Gordon Hirabayashi, for violating the curfew. Military leaders went on to enforce Roosevelt's order aggressively by requiring that persons of Japanese ancestry living on the West Coast be sent to internment camps.

Fred Korematsu was a native-born citizen of the United States of Japanese ancestry. After Pearl Harbor he volunteered to join the army but was rejected due to poor health. He then worked for the civilian defense industry as a welder. Korematsu was convicted of refusing to report to the authorities for eventual internment in a camp. He appealed to the Supreme Court, which after bitter internal disputes refused to overturn the military's judgment and allowed the camps. Some justices wanted to allow only the curfew but not the camps; others refused to second-guess the president, Congress, and the military at all. One of the eventual dissenters, Justice Murphy, after reading a draft of an opinion from another potential dissenting justice, sent a copy of the draft to his clerk with the following note: "Gene: Read this and perish! The Court has blown up over the Jap case—just as I expected it would."

This case set the stage for future ones involving "Equal Protection" by introducing the concepts of "suspect" classification and "strict scrutiny."

Mr. Justice Black delivered the opinion of the Court: It should be noted, to begin with, that all legal restrictions which curtail the civil rights of a single racial group are immediately suspect. That is not to say that all such restrictions are unconstitutional. It is to say that courts must subject them to the most rigid scrutiny. Pressing public necessity may sometimes justify the existence of such restrictions; racial antagonism never can. . . .

In the light of the principles we announced in the *Hirabayashi* case, we are unable to conclude that it was beyond the war power of Congress and the Executive to exclude those of Japanese ancestry from the West Coast war area at the time they did. True, exclusion from the area in which one's home is located is a far greater deprivation than constant confinement to the home from 8:00 p.m. to 6:00 a.m. Nothing short of apprehension by the proper military authorities of the gravest imminent danger to the public safety can constitutionally justify either. But exclusion from a threatened area, no less than curfew, has a definite and close relationship to the prevention of espionage and sabotage. The military authorities, charged with the primary responsibility of defending our shores, concluded that curfew provided inadequate protection and ordered exclusion. They did so, as pointed out in our *Hirabayashi* opinion, in accordance with Congressional authority to the military to say who should, and who should not, remain in the threatened areas. . . .

Like curfew, exclusion of those of Japanese origin was deemed necessary because of the presence of an unascertained number of disloyal members of the group, most of whom we have no doubt were loyal to this country. It was because we could not reject the finding of the military authorities that it was impossible to bring about an immediate segregation of the disloyal from the loyal that we sustained the validity of the curfew order as applying to the whole group. In the instant case, temporary exclusion of the entire group was rested by the military on the same ground. The judgment that exclusion of the whole group was for the same reason a military imperative answers the contention that the exclusion was in the nature of group punishment based on antagonism to those of Japanese origin. That there were members of the group who retained loyalties to Japan has been confirmed by investigations made subsequent to the exclusion. Approximately five thousand American citizens of Japanese ancestry refused to swear unqualified allegiance to the United States and to renounce allegiance to the Japanese Emperor, and several thousand evacuees requested repatriation to Japan.

We uphold the exclusion order as of the time it was made and when the petitioner violated it. . . . In doing so, we are not unmindful of the hardships imposed by it upon a large group of American citizens. . . . But hardships are part of

war, and war is an aggregation of hardships. All citizens alike, both in and out of uniform, feel the impact of war in greater or lesser measure. Citizenship has its responsibilities as well as its privileges, and in time of war the burden is always heavier. Compulsory exclusion of large groups of citizens from their homes, except under circumstances of direst emergency and peril, is inconsistent with our basic governmental institutions. But when under conditions of modern warfare our shores are threatened by hostile forces, the power to protect must be commensurate with the threatened danger. . . .

It is said that we are dealing here with the case of imprisonment of a citizen in a concentration camp solely because of his ancestry, without evidence or inquiry concerning his loyalty and good disposition towards the United States. Our task would be simple, our duty clear, were this a case involving the imprisonment of a loyal citizen in a concentration camp because of racial prejudice. Regardless of the true nature of the assembly and relocation centers—and we deem it unjustifiable to call them concentration camps with all the ugly connotations that term implies—we are dealing specifically with nothing but an exclusion order. To cast this case into outlines of racial prejudice, without reference to the real military dangers which were presented, merely confuses the issue. Korematsu was not excluded from the Military Areas because of hostility to him or his race. He *was* excluded because we are at war with the Japanese Empire, because the properly constituted military authorities feared an invasion of our West Coast and felt constrained to take proper security measures, because they decided that the military urgency of the situation demanded that all citizens of Japanese ancestry be segregated from the West Coast temporarily, and finally, because Congress, reposing its confidence in this time of war in our military leaders—as inevitably it must—determined that they should have the power to do just this. There was evidence of disloyalty on the part of some, the military authorities considered that the need for action was great, and time was short. We cannot—by availing ourselves of the calm perspective of hindsight—now say that at that time these actions were unjustified.

Affirmed.

Mr. Justice Murphy, dissenting: This exclusion of "all persons of Japanese ancestry, both alien and non-alien," from the Pacific Coast area on a plea of military necessity in the absence of martial law ought not to be approved. Such exclusion goes over "the very brink of constitutional power" and falls into the ugly abyss of racism.

In dealing with matters relating to the prosecution and progress of a war, we must accord great respect and consideration to the judgments of the military authorities who are on the scene and who have full knowledge of the military facts. . . .

At the same time, however, it is essential that there be definite limits to military discretion, especially where martial law has not been declared.

That this forced exclusion was the result in good measure of this erroneous assumption of racial guilt rather than bona fide military necessity is evidenced by the Commanding General's Final Report on the evacuation. . . . In it he refers to all individuals of Japanese descent as "subversive," as belonging to "an enemy race" whose "racial strains are undiluted," and as constituting "over 112,000 potential enemies . . . at large today." . . . In support of this blanket condemnation of all persons of Japanese descent, however, no reliable evidence is cited to show that such individuals were generally disloyal, or . . . constitute[d] a special menace to defense installations or war industries, or had otherwise by their behavior furnished reasonable ground for their exclusion as a group.

Justification for the exclusion is sought, instead, mainly upon questionable racial and sociological grounds not ordinarily within the realm of expert military judgment, supplemented by certain semi-military conclusions drawn from an unwarranted use of circumstantial evidence. Individuals of Japanese ancestry are condemned because they are said to be "a large, unassimilated, tightly knit racial group, bound to an enemy nation by strong ties of race, culture, custom and religion." They are claimed to be given to "emperor worshipping ceremonies" and to "dual citizenship." Japanese language schools and allegedly pro-Japanese organizations are cited as evidence of possible group disloyalty, together with facts as to certain persons being educated and residing at length in Japan. . . .

The main reasons relied upon by those responsible for the forced evacuation . . . [are] largely an

accumulation of much of the misinformation, half-truths and insinuations that for years have been directed against Japanese Americans by people with racial and economic prejudices—the same people who have been among the foremost advocates of the evacuation. A military judgment based upon such racial and sociological considerations is not entitled to the great weight ordinarily given the judgments based upon strictly military considerations. Especially is this so when every charge relative to race, religion, culture, geographical location, and legal and economic status has been substantially discredited by independent studies made by experts in these matters.

No adequate reason is given for the failure to treat these Japanese Americans on an individual basis by holding investigations and hearings to separate the loyal from the disloyal, as was done in the case of persons of German and Italian ancestry. It is asserted merely that the loyalties of this group "were unknown and time was of the essence." Yet nearly four months elapsed after Pearl Harbor before the first exclusion order was issued; nearly eight months went by until the last order was issued; and the last of these "subversive" persons was not actually removed until almost eleven months had elapsed. Leisure and deliberation seem to have been more of the essence than speed. And the fact that conditions were not such as to warrant a declaration of martial law adds strength to the belief that the factors of time and military necessity were not as urgent as they have been represented to be.

Moreover, there was no adequate proof that the Federal Bureau of Investigation and the military and naval intelligence services did not have the espionage and sabotage situation well in hand during this long period. Nor is there any denial of the fact that not one person of Japanese ancestry was accused or convicted of espionage or sabotage after Pearl Harbor while they were still free, a fact which is some evidence of the loyalty of the vast majority of those individuals and of the effectiveness of the established methods of combating these evils. It seems incredible that under these circumstances it would have been impossible to hold loyalty hearings for the mere 112,000 persons involved—or at least for the 70,000 American citizens—especially when a large part of this number represented children and elderly men and women. . . .

I dissent, therefore, from this legalization of racism. Racial discrimination in any form and in any degree has no justifiable part whatever in our democratic way of life. It is unattractive in any setting but it is utterly revolting among a free people who have embraced the principles set forth in the Constitution of the United States. All residents of this nation are kin in some way by blood or culture to a foreign land. Yet they are primarily and necessarily a part of the new and distinct civilization of the United States. They must accordingly be treated at all times as the heirs of the American experiment and as entitled to all the rights and freedoms guaranteed by the Constitution.

Mr. Justice Jackson, dissenting: Korematsu was born on our soil, of parents born in Japan. The Constitution makes him a citizen of the United States by nativity and a citizen of California by residence. No claim is made that he is not loyal to this country. There is no suggestion that apart from the matter involved here he is not law-abiding and well disposed. Korematsu, however, has been convicted of an act not commonly a crime. It consists merely of being present in the state whereof he is a citizen, near the place where he was born, and where all his life he has lived.

Even more unusual is the series of military orders which made this conduct a crime. They forbid such a one to remain, and they also forbid him to leave. They were so drawn that the only way Korematsu could avoid violation was to give himself up to the military authority. This meant submission to custody, examination, and transportation out of the territory, to be followed by indeterminate confinement in detention camps.

A citizen's presence in the locality, however, was made a crime only if his parents were of Japanese birth. Had Korematsu been one of four—the others being, say, a German alien enemy, an Italian alien enemy, and a citizen of American-born ancestors, convicted of treason but out on parole—only Korematsu's presence would have violated the order. The difference between their innocence and his crime would result, not from anything he did, said or thought, different than they, but only in that he was born of different racial stock.

Now, if any fundamental assumption underlies our system, it is that guilt is personal and not inheritable. Even if all of one's antecedents had been convicted of treason, the Constitution forbids its penalties to be visited upon him, for it provides that "no attainder of treason shall work corruption of blood, or forfeiture except during the life of the person attainted." But here is an attempt to . . . make an otherwise innocent act a crime merely because this prisoner is the son of parents as to whom he had no choice, and belongs to a race from which there is no way to resign.

. . . [T]he "law" which this prisoner is convicted of disregarding is not found in an act of Congress, but in a military order. Neither the Act of Congress nor the Executive Order of the President, nor both together, would afford a basis for the conviction. . . . And it is said that if the military commander had reasonable military grounds for promulgating the orders, they are constitutional and become law, and the Court is required to enforce them. . . .

It would be impracticable and dangerous idealism to expect or insist that each specific military command in an area of probable operations will conform to conventional tests of constitutionality . . . The armed services must protect a society, not merely its Constitution. The very essence of the military job is to marshal physical force, to remove every obstacle to its effectiveness, to give it every strategic advantage. Defense measures will not, and often should not, be held within the limits that bind civil authority in peace. . . .

But if we cannot confine military expedients by the Constitution, neither would I distort the Constitution to approve all that the military may deem expedient. . . . I cannot say . . . that the orders of General DeWitt were not reasonably expedient military precautions, nor could I say that they were. But even if they were permissible military procedures, I deny that it follows that they are constitutional. If, as the Court holds, it does follow, then we may as well say that any military order will be constitutional and have done with it. . . .

Much is said of the danger to liberty from the Army program. . . . But a judicial construction of the due process clause that will sustain this order is a far more subtle blow to liberty. . . . A military order, however unconstitutional, is not apt to last longer than the military emergency. Even during that period a succeeding commander may revoke it all. But once a judicial opinion rationalizes such an order to show that it conforms to the Constitution, or rather rationalizes the Constitution to show that the Constitution sanctions such an order, the Court for all time has validated the principle of racial discrimination in criminal procedure and of transplanting American citizens. The principle then lies about like a loaded weapon ready for the hand of any authority that can bring forward a plausible claim of an urgent need. . . . All who observe the work of courts are familiar with what Judge Cardozo described as "the tendency of a principle to expand itself to the limit of its logic." A military commander may overstep the bounds of constitutionality, and it is an incident. But if we review and approve, that passing incident becomes the doctrine of the Constitution. There it has a generative power of its own, and all that it creates will be in its own image. Nothing better illustrates this danger than does the Court's opinion in this case.

It argues that we are bound to uphold the conviction of Korematsu because we upheld one in *Hirabayashi*. . . .

In that case we were urged to consider only the curfew feature. . . . We yielded, and the Chief Justice guarded the opinion as carefully as language will do. He said: ". . . We decide only the issue as we have defined it—we decide only that the *curfew order* as applied, and at the time it was applied, was within the boundaries of the war power." . . . Now the principle of racial discrimination is pushed from support of mild measures to very harsh ones, and from temporary deprivations to indeterminate ones. And the precedent which it is said requires us to do so is *Hirabayashi.* The Court is now saying that in *Hirabayashi* we did decide the very things we there said we were not deciding. Because we said that these citizens could be made to stay in their homes during the hours of dark, it is said we must require them to leave home entirely; and if that, we are told they may also be taken into custody for deportation; and if that, it is argued they may also be held for some undetermined time in detention camps. How far the principle of this case would be extended before plausible reasons would play out, I do not know.

I should hold that a civil court cannot be made to enforce an order which violates constitutional

limitations even if it is a reasonable exercise of military authority. The courts can exercise only the judicial power, can apply only law, and must abide by the Constitution, or they cease to be civil courts and become instruments of military policy.

. . . I would not lead people to rely on this Court for a review that seems to me wholly delusive. . . . If the people ever let command of the war power fall into irresponsible and unscrupulous hands, the courts wield no power equal to its restraint. The chief restraint upon those who command the physical forces of the country . . . must be their responsibility to the political judgments of their contemporaries and to the moral judgments of history.

My duties as a justice . . . do not require me to make a military judgment as to whether General DeWitt's evacuation and detention program was a reasonable military necessity. I do not suggest that the courts should have attempted to interfere with the Army in carrying out its task. But I do not think they may be asked to execute a military expedient that has no place in law under the Constitution. I would reverse the judgment and discharge the prisoner.

REVIEW AND DISCUSSION QUESTIONS

1. Explain the reasoning behind Justice Black's opinion for the Court.
2. Justice Black begins his opinion with a brief indication of how he approaches the case, mentioning both suspect restrictions and rigid (or strict) scrutiny. These concepts have now become the centerpiece of Equal Protection analysis. Explain what he means.
3. Describe the basis of Justice Murphy's dissent.
4. Justice Jackson's dissent focuses on due process rather than equal protection. Explain how that aspect of Justice Jackson's position differs from Justice Murphy's.
5. On whose shoulders does Justice Jackson think rests the possibility of avoiding future abuses of authority by the military? Explain.
6. Justice Jackson's conclusion has been described as an example of judicial "nihilism." Explain that comment, indicating if you agree.
7. In later years Justice Douglas, who had reluctantly gone along with the majority, came to regret his decision and believed *Korematsu* was a grave mistake. He also said, however, that he thought any of the many courts he served on during his long career would have ruled the same way. Justice Black, on the other hand, continued to believe that Congress had not exceeded its power in delegating authority to the military to intern Japanese Americans. Many today think the case shows how willing the Court and Congress are to bend the Constitution in time of war. Still others see it as an indication that the racism infecting society has not been limited to blacks. How would you assess the case?

Racial Equality and Affirmative Action

Ronald Dworkin

The words "affirmative action" were first used by President John Kennedy in 1961. His Executive Order 10925 required that contractors working for the federal government take "affirmative action to ensure that applicants are employed without regard to their race, creed, color, or national origin." Over the next decade, however, this requirement moved slowly away from racial neutrality toward preferences and quotas, as the government tried to expand opportunities for minorities. In 1965 President Lyndon Johnson's secretary of labor issued guidelines for hiring federal contractors that included "goals

and timetables" to achieve increased minority employment. As originally understood, affirmative action is largely uncontroversial, requiring only that recruitment of minorities be seriously undertaken and that hiring should be done without regard to race and strictly on the basis of ability to perform. Now, however, affirmative action means more than that: it includes special minority scholarships; preferences for minorities in terms of jobs, contracts, and admissions; and sometimes implicit quotas in hiring and admission. The most important of the Supreme Court cases to consider affirmative action is *Regents of the University of California v. Bakke* (1978).

Allan Bakke was turned down for admission to the University of California at Davis medical school, which had set aside sixteen of its hundred slots for "disadvantaged" students. The evidence at his trial showed that although in theory whites could be admitted under the affirmative action program, none in fact were. It also showed that Bakke's overall grade-point average was 3.46, while the average for the affirmative action program was 2.62. His Medical College Aptitude Tests percentile scores were verbal, 96; quantitative, 94; and science, 97. These compared with 34, 30, and 37, respectively, for the affirmative action admittees. Bakke also received an overall rating of 94 out of possible 100 in his interview. As to the regular admission process, that year there were one African American, six Mexican Americans, and thirty-seven Asian Americans admitted.

In this essay, written before the Supreme Court handed down its decision in the case, Ronald Dworkin examines the different arguments open to Bakke's lawyer and contends that the Court should rule against Bakke.

On October 12, 1977 the Supreme Court heard oral argument in the case of *The Regents of the University of California v. Allan Bakke.* No lawsuit has ever been more widely watched or more thoroughly debated in the national and international press before the Court's decision. Still, some of the most pertinent facts set before the Court have not been clearly summarized.

The medical school of the University of California at Davis has an affirmative action program (called the "task force program") designed to admit more black and other minority students. It sets sixteen places aside for which only members of "educationally and economically disadvantaged minorities" compete. Allan Bakke, white, applied for one of the remaining eighty-four places; he was rejected but, since his test scores were relatively high, the medical school has conceded that it could not prove that he would have been rejected if the sixteen places reserved had been open to him. Bakke sued, arguing that the task force program deprived him of his constitutional rights. The California Supreme Court agreed, and ordered the medical school to admit him. The university appealed to the Supreme Court.

The Davis program for minorities is in certain respects more forthright (some would say cruder) than similar plans now in force in many other American universities and professional schools. Such programs aim to increase the enrollment of black and other minority students by allowing the fact of their race to count affirmatively as part of the case for admitting them. Some schools set a "target" of a particular number of minority places instead of setting aside a flat number of places. But Davis would not fill the number of places set aside unless there were sixteen minority candidates it considered clearly qualified for medical education. The difference is therefore one of administrative strategy and not of principle.

So the constitutional question raised by *Bakke* is of capital importance for higher education in America, and a large number of universities and schools have entered briefs *amicus curiae* urging the Court to reverse the California decision. They believe that if the decision is affirmed then they will no longer be free to use explicit racial criteria in any part of their admissions programs, and that they will therefore be unable to fulfill what they take to be their responsibilities to the nation.

It is often said that affirmative action programs aim to achieve a racially conscious society divided into racial and ethnic groups, each entitled, as a group, to some proportionable share of resources, careers, or opportunities. That is a perverse description. American society is currently a racially conscious society; that is the inevitable and evident consequence of a history of slavery, repression, and

prejudice. Black men and women, boys and girls, are not free to choose for themselves in what roles—or as members of which social groups—others will characterize them. They are black, and no other feature of personality or allegiance or ambition will so thoroughly influence how they will be perceived and treated by others, and the range and character of the lives that will be open to them.

The tiny number of black doctors and professionals is both a consequence and a continuing cause of American racial consciousness, one link in a long and self-fueling chain reaction. Affirmative action programs use racially explicit criteria because their immediate goal is to increase the number of members of certain races in these professions. But their long-term goal is to *reduce* the degree to which American society is overall a racially conscious society.

The programs rest on two judgments. The first is a judgment of social theory: that America will continue to be pervaded by racial divisions as long as the most lucrative, satisfying, and important careers remain mainly the prerogative of members of the white race, while others feel themselves systematically excluded from a professional and social elite. The second is a calculation of strategy: that increasing the number of blacks who are at work in the professions will, in the long run, reduce the sense of frustration and injustice and racial self-consciousness in the black community to the point at which blacks may begin to think of themselves as individuals who can succeed like others through talent and initiative. At that future point the consequences of nonracial admissions programs, whatever these consequences might be, could be accepted with no sense of racial barriers or injustice.

It is therefore the worst possible misunderstanding to suppose that affirmative action programs are designed to produce a balkanized America, divided into racial and ethnic subnations. They use strong measures because weaker ones will fail; but their ultimate goal is to lessen not to increase the importance of race in American social and professional life.

According to the 1970 census, only 2.1 percent of US doctors were black. Affirmative action programs aim to provide more black doctors to serve black patients. This is not because it is desirable that blacks treat blacks and whites treat whites, but because blacks, for no fault of their own, are now unlikely to be well served by whites, and because a failure to provide the doctors they trust will exacerbate rather than reduce the resentment that now leads them to trust only their own. Affirmative action tries to provide more blacks as classmates for white doctors, not because it is desirable that a medical school class reflect the racial makeup of the community as a whole, but because professional association between blacks and whites will decrease the degree to which whites think of blacks as a race rather than as people, and thus the degree to which blacks think of themselves that way. It tries to provide "role models" for future black doctors, not because it is desirable for a black boy or girl to find adult models only among blacks, but because our history has made them so conscious of their race that the success of whites, for now, is likely to mean little or nothing for them.

The history of the campaign against racial injustice since 1954, when the Supreme Court decided *Brown v. Board of Education,* is a history in large part of failure. We have not succeeded in reforming the racial consciousness of our society by racially neutral means. We are therefore obliged to look upon the arguments for affirmative action with sympathy and an open mind. Of course, if Bakke is right that such programs, no matter how effective they may be, violate his constitutional rights then they cannot be permitted to continue. But we must not forbid them in the name of some mindless maxim, like the maxim that it cannot be right to fight fire with fire, or that the end cannot justify the means. If the strategic claims for affirmative action are cogent, they cannot be dismissed simply on the ground that racially explicit tests are distasteful. If such tests are distasteful it can only be for reasons that make the underlying social realities the programs attack more distasteful still.

The New Republic, in a recent editorial opposing affirmative action, missed that point. "It is critical to the success of a liberal pluralism," it said, "that group membership itself is not among the permissible criteria of inclusion and exclusion." But group membership is in fact, as a matter of social reality rather than formal admission standards, part of what determines inclusion or exclusion for us now. If we must choose between a

society that is in fact liberal and an illiberal society that scrupulously avoids formal racial criteria, we can hardly appeal to the ideals of liberal pluralism to prefer the latter.

Professor Archibald Cox of Harvard Law School, speaking for the University of California in oral argument, told the Supreme Court that this is the choice the United States must make. As things stand, he said, affirmative action programs are the only effective means of increasing the absurdly small number of black doctors. The California Supreme Court, in approving Bakke's claim, had urged the university to pursue that goal by methods that do not explicitly take race into account. But that is unrealistic. We must distinguish, as Cox said, between two interpretations of what the California court's recommendation means. It might mean that the university should aim at the same immediate goal, of increasing the proportion of black and other minority students in the medical school, by an admissions procedure that on the surface is not racially conscious.

That is a recommendation of hypocrisy. If those who administer the admissions standards, however these are phrased, understand that their immediate goal is to increase the number of blacks in the school, then they will use race as a criterion in making the various subjective judgments the explicit criteria will require, because that will be, given the goal, the only right way to make those judgments. The recommendation might mean, on the other hand, that the school should adopt some nonracially conscious goal, like increasing the number of disadvantaged students of all races, and then hope that that goal will produce an increase in the number of blacks as a by-product. But even if that strategy is less hypocritical (which is far from plain), it will almost certainly fail because no different goal, scrupulously administered in a nonracially conscious way, will in fact significantly increase the number of black medical students.

Cox offered powerful evidence for that conclusion, and it is supported by the recent and comprehensive report of the Carnegie Council on Policy Studies in Higher Education. Suppose, for example, that the medical school sets aside separate places for applicants "disadvantaged" on some racially neutral test, like poverty, allowing only those disadvantaged in that way to compete for these places. If the school selects these from that group who scored best on standard medical school aptitude tests, then it will take almost no blacks, because blacks score relatively low even among the economically disadvantaged. But if the school chooses among the disadvantaged on some basis other than test scores, just so that more blacks will succeed, then it will not be administering the special procedure in a nonracially conscious way.

So Cox was able to put his case in the form of two simple propositions. A racially conscious test for admission, even one that sets aside certain places for qualified minority applicants exclusively, serves goals that are in themselves unobjectionable and even urgent. Such programs are, moreover, the only means that offer any significant promise of achieving these goals. If these programs are halted, then no more than a trickle of black students will enter medical or other professional schools for another generation at least.

If these propositions are sound, then on what ground can it be thought that such programs are either wrong or unconstitutional? We must notice an important distinction between two different sorts of objections that might be made. These programs are intended, as I said, to decrease the importance of race in the United States in the long run. It may be objected, first, that the programs will in fact harm that goal more than they will advance it. There is no way now to prove that that is so. Cox conceded, in his argument, that there are costs and risks in these programs.

Affirmative action programs seem to encourage, for example, a popular misunderstanding, which is that they assume that racial or ethnic groups are entitled to proportionate shares of opportunities, so that Italian or Polish ethnic minorities are, in theory, as entitled to their proportionate shares as blacks or Chicanos or American Indians are entitled to the shares the present programs give them. That is a plain mistake: the programs are not based on the idea that those who are aided are entitled to aid, but only on the strategic hypothesis that helping them is now an effective way of attacking a national problem. Some medical schools may well make that judgment, under certain circumstances, about a white ethnic minority. Indeed it seems likely that some

medical schools are even now attempting to help white Appalachian applicants, for example, under programs of regional distribution.

So the popular understanding is wrong, but so long as it persists it is a cost of the program because the attitudes it encourages tend to a degree to make people more rather than less conscious of race. There are other possible costs. It is said, for example, that some blacks find affirmative action degrading; they find that it makes them more rather than less conscious of prejudice against their race as such. This attitude is also based on a misperception, I think, but for a small minority of blacks at least it is a genuine cost.

In the view of the many important universities who have such programs, however, the gains will very probably exceed the losses in reducing racial consciousness overall. This view is hardly so implausible that it is wrong for these universities to seek to acquire the experience that will allow us to judge whether they are right. It would be particularly silly to forbid these experiments if we know that the failure to try will mean, as the evidence shows, that the status quo will almost certainly continue. In any case, this first objection could provide no argument that would justify a decision by the Supreme Court holding the programs unconstitutional. The Court has no business substituting its speculative judgment about the probable consequences of educational policies for the judgment of professional educators.

So the acknowledged uncertainties about the long-term results of such programs could not justify a Supreme Court decision making them illegal. But there is a second and very different form of objection. It may be argued that even if the programs *are* effective in making our society less a society dominated by race, they are nevertheless unconstitutional because they violate the individual constitutional rights of those, like Allan Bakke, who lose places in consequence. In the oral argument Reynold H. Colvin of San Francisco, who is Bakke's lawyer, made plain that his objection takes this second form. Mr. Justice White asked him whether he accepted that the goals affirmative action programs seek are important goals. Mr. Colvin acknowledged that they were. Suppose, Justice White continued, that affirmative action programs are, as Cox had argued, the only effective means of seeking such goals. Would Mr. Colvin nevertheless maintain that the programs are unconstitutional? Yes, he insisted, they would be, because his client has a constitutional right that the programs be abandoned, no matter what the consequences.

Mr. Colvin was wise to put his objections on this second ground; he was wise to claim that his client has rights that do not depend on any judgment about the likely consequences of affirmative action for society as a whole, because if he makes out that claim then the Court must give him the relief he seeks.

But can he be right? If Allan Bakke has a constitutional right so important that the urgent goals of affirmative action must yield, then this must be because affirmative action violates some fundamental principle of political morality. This is not a case in which what might be called formal or technical law requires a decision one way or the other. There is no language in the Constitution whose plain meaning forbids affirmative action. Only the most naïve theories of statutory construction could argue that such a result is required by the language of any earlier Supreme Court decision or of the Civil Rights Act of 1964 or of any other congressional enactment. If Mr. Colvin is right it must be because Allan Bakke has not simply some technical legal right but an important moral right as well.

What could that right be? The popular argument frequently made on editorial pages is that Bakke has a right to be judged on his merit. Or that he has a right to be judged as an individual rather than as a member of a social group. Or that he has a right, as much as any black man, not to be sacrificed or excluded from any opportunity because of his race alone. But these catch phrases are deceptive here, because, as reflection demonstrates, the only genuine principle they describe is the principle that no one should suffer from the prejudice or contempt of others. And that principle is not at stake in this case at all. In spite of popular opinion, the idea that the *Bakke* case presents a conflict between a desirable social goal and important individual rights is a piece of intellectual confusion.

Consider, for example, the claim that individuals applying for places in medical school should be judged on merit, and merit alone. If that slogan means that admissions committees should take nothing into account but scores on some particular intelligence test, then it is arbitrary and, in any

case, contradicted by the long-standing practice of every medical school. If it means, on the other hand, that a medical school should choose candidates that it supposes will make the most useful doctors, then everything turns on the judgment of what factors make different doctors useful. The Davis medical school assigned to each regular applicant, as well as to each minority applicant, what it called a "benchmark score." This reflected not only the results of aptitude tests and college grade averages, but a subjective evaluation of the applicant's chances of functioning as an effective doctor, in view of society's present needs for medical service. Presumably the qualities deemed important were different from the qualities that a law school or engineering school or business school would seek, just as the intelligence tests a medical school might use would be different from the tests these other schools would find appropriate.

There is no combination of abilities and skills and traits that constitutes "merit" in the abstract; if quick hands count as "merit" in the case of a prospective surgeon, this is because quick hands will enable him to serve the public better and for no other reason. If a black skin will, as a matter of regrettable fact, enable another doctor to do a different medical job better, then that black skin is by the same token "merit" as well. That argument may strike some as dangerous; but only because they confuse its conclusion—that black skin may be a socially useful trait in particular circumstances—with the very different and despicable idea that one race may be inherently more worthy than another.

Consider the second of the catch phrases I have mentioned. It is said that Bakke has a right to be judged as an "individual," in deciding whether he is to be admitted to medical school and thus to the medical profession, and not as a member of some group that is being judged as a whole. What can that mean? Any admissions procedure must rely on generalizations about groups that are justified only statistically. The regular admissions process at Davis, for example, set a cutoff figure for college grade-point averages. Applicants whose averages fell below that figure were not invited to any interview, and therefore rejected out of hand.

An applicant whose average fell one point below the cutoff might well have had personal qualities of dedication or sympathy that would have been revealed at an interview, and that would have made him or her a better doctor than some applicant whose average rose one point above the line. But the former is excluded from the process on the basis of a decision taken for administrative convenience and grounded in the generalization, unlikely to hold true for every individual, that those with grade averages below the cutoff will not have other qualities sufficiently persuasive. Indeed, even the use of standard Medical College Aptitude Tests (MCAT) as part of the admissions procedure requires judging people as part of groups because it assumes that test scores are a guide to medical intelligence which is in turn a guide to medical ability. Though this judgment is no doubt true statistically, it hardly holds true for every individual.

Allan Bakke was himself refused admission to two other medical schools, not because of his race but because of his age: these schools thought that a student entering medical school at the age of thirty-three was likely to make less of a contribution to medical care over his career than someone entering at the standard age of twenty-one. Suppose these schools relied, not on any detailed investigation of whether Bakke himself had abilities that would contradict the generalization in his specific case, but on a rule of thumb that allowed only the most cursory look at applicants over (say) the age of thirty. Did these two medical schools violate his right to be judged as an individual rather than as a member of a group?

The Davis Medical School permitted whites to apply for the sixteen places reserved for members of "educationally or economically disadvantaged minorities," a phrase whose meaning might well include white ethnic minorities. In fact several whites have applied, though none has been accepted, and the California Court found that the special committee charged with administering the program had decided, in advance, against admitting any. Suppose that decision had been based on the following administrative theory: it is so unlikely that any white doctor can do as much to counteract racial imbalance in the medical professions as a well-qualified and trained black doctor can do that the committee should for reasons of convenience proceed on the presumption no white doctor could.

That presumption is, as a matter of fact, more plausible than the corresponding presumption about medical students over the age of thirty, or even the presumption about applicants whose grade-point averages fall below the cutoff line. If the latter presumptions do not deny the alleged right of individuals to be judged as individuals in an admissions procedure, then neither can the former.

Mr. Colvin, in oral argument, argued the third of the catch phrases I mentioned. He said that his client had a right not to be excluded from medical school because of his race alone, and this as a statement of constitutional right sounds more plausible than claims about the right to be judged on merit or as an individual. It sounds plausible, however, because it suggests the following more complex principle. Every citizen has a constitutional right that he not suffer disadvantage, at least in the competition for any public benefit, because the race or religion or sect or region or other natural or artificial group to which he belongs is the object of prejudice or contempt.

That is a fundamentally important constitutional right, and it is that right that was systematically violated for many years by racist exclusions and anti-Semitic quotas. Color bars and Jewish quotas were not unfair just because they made race or religion relevant or because they fixed on qualities beyond individual control. It is true that blacks or Jews do not choose to be blacks or Jews. But it is also true that those who score low in aptitude or admissions tests do not choose their levels of intelligence. Nor do those denied admission because they are too old, or because they do not come from a part of the country underrepresented in the school, or because they cannot play basketball well, choose not to have the qualities that made the difference.

Race seems different because exclusions based on race have historically been motivated not by some instrumental calculation, as in the case of intelligence or age or regional distribution or athletic ability, but because of contempt for the excluded race or religion as such. Exclusion by race was in itself an insult, because it was generated by and signaled contempt.

Bakke's claim, therefore, must be made more specific than it is. He says he was kept out of medical school because of his race. Does he mean that he was kept out because his race is the object of prejudice or contempt? That suggestion is absurd. A very high proportion of those who were accepted (and, presumably, of those who run the admissions program) were members of the same race. He therefore means simply that if he had been black he would have been accepted, with no suggestion that this would have been so because blacks are thought more worthy or honorable than whites.

That is true: no doubt he would have been accepted if he were black. But it is also true, and in exactly the same sense, that he would have been accepted if he had been more intelligent, or made a better impression in his interview, or, in the case of other schools, if he had been younger when he decided to become a doctor. Race is not, in, *his* case, a different matter from these other factors equally beyond his control. It is not a different matter because in his case race is not distinguished by the special character of public insult. On the contrary the program presupposes that his race is still widely if wrongly thought to be superior to others.

In the past, it made sense to say that an excluded black or Jewish student was being sacrificed because of his race or religion; that meant that his or her exclusion was treated as desirable in itself, not because it contributed to any goal in which he as well as the rest of society might take pride. Allan Bakke is being "sacrificed" because of his race only in a very artificial sense of the word. He is being "sacrificed" in the same artificial sense because of his level of intelligence, since he would have been accepted if he were more clever than he is. In both cases he is being excluded not by prejudice but because of a rational calculation about the socially most beneficial use of limited resources for medical education.

It may now be said that this distinction is too subtle, and that if racial classifications have been and may still be used for malign purposes, then everyone has a flat right that racial classifications not be used at all. This is the familiar appeal to the lazy virtue of simplicity. It supposes that if a line is difficult to draw, or might be difficult to administer if drawn, then there is wisdom in not making the attempt to draw it. There may be cases in which that is wise, but those would be cases in which nothing of great value would as a consequence be

lost. If racially conscious admissions policies now offer the only substantial hope for bringing more qualified black and other minority doctors into the profession, then a great loss is suffered if medical schools are not allowed voluntarily to pursue such programs. We should then be trading away a chance to attack certain and present injustice in order to gain protection we may not need against speculative abuses we have other means to prevent. And such abuses cannot, in any case, be worse than the injustice to which we would then surrender.

We have now considered three familiar slogans, each widely thought to name a constitutional right that enables Allan Bakke to stop programs of affirmative action no matter how effective or necessary these might be. When we inspect these slogans, we find that they can stand for no genuine principle except one. This is the important principle that no one in our society should suffer because he is a member of a group thought less worthy of respect, as a group, than other groups. We have different aspects of that principle in mind when we say that individuals should be judged on merit, that they should be judged as individuals, and that they should not suffer disadvantages because of their race. The spirit of that fundamental principle is the spirit of the goal that affirmative action is intended to serve. The principle furnishes no support for those who find, as Bakke does, that their own interests conflict with that goal.

It is of course regrettable when any citizen's expectations are defeated by new programs serving some more general concern. It is regrettable, for example, when established small businesses fail because new and superior roads are built; in that case people have invested more than Bakke has. And they have more reason to believe their businesses will continue than Bakke had to suppose he could have entered the Davis medical school at thirty-three even without a task force program.

There is, of course, no suggestion in the program that Bakke shares in any collective or individual guilt for racial injustice in America; or that he is any less entitled to concern or respect than any black student accepted in the program. He has been disappointed, and he must have the sympathy due that disappointment, just as any other disappointed applicant—even one with much worse test scores who would not have been accepted in any event—must have sympathy. Each is disappointed because places in medical schools are scarce resources and must be used to provide what the more general society most needs. It is hardly Bakke's fault that racial justice is now a special need—but he has no right to prevent the most effective measures of securing that justice from being used.

REVIEW AND DISCUSSION QUESTIONS

1. Dworkin says affirmative action programs rest on two premises. Explain them.
2. How does Dworkin answer critics of those two goals?
3. How does Dworkin respond to the claim that Bakke's right to be judged on the basis of "merit" was infringed?
4. How does he respond to the claim that Bakke was not "judged as an individual"?
5. In fact Bakke's lawyer claimed he had a right not to be excluded from medical school based on race. How does Dworkin answer that charge?
6. How might it be argued that some people have a *right* to affirmative action programs? Are those groups the same ones that benefit from such a program? Explain.
7. How might it be argued that affirmative action programs actually harm relations between minorities and others? How might it help?
8. "Bakke had more 'merit' as a future student, but not as a prospective doctor." Discuss. What, exactly, is merit?

Affirmative Action in Universities

Grutter v. Bollinger and *Gratz v. Bollinger*

In June 2003, in two cases involving the University of Michigan, the Supreme Court took up the issue of affirmative action in higher education for the first time since the *Bakke* case of 1978. In *Grutter* v. *Bollinger*, the Court held that the University of Michigan Law School's use of race as a factor in student admission is constitutionally permissible. In delivering the majority opinion, Justice O'Connor explains that the law school has a compelling interest in attaining a diverse student body and that its flexible, individualized admissions process is narrowly tailored to further that interest. However, in a vigorous dissent, Justice Thomas responds that marginal improvements in legal education cannot justify racial discrimination, and he criticizes the law school for not seeking race-neutral ways to enhance diversity.

On the same day, the Court ruled in a companion case, *Gratz* v. *Bollinger*, that the University's use of race in its undergraduate admissions process was constitutionally flawed because it guaranteed members of underrepresented minorities 20 points on a 150-point scale. In their dissents to this decision, Justice Souter argues that the undergraduate admissions policy simply does with a numbered scale what the law school accomplishes with its "holistic" review, while Justice Ginsburg stresses the continuing legacy "of a system of racial caste only recently ended" and the difference between using race to exclude and taking race into account for the purpose of promoting equality.

Grutter v. Bollinger

Justice O'Connor delivered the opinion of the Court: This case requires us to decide whether the use of race as a factor in student admissions by the University of Michigan Law School (Law School) is unlawful.

I

The Law School ranks among the Nation's top law schools. . . . The Law School looks for individuals with "substantial promise for success in law school" and "a strong likelihood of succeeding in the practice of law and contributing in diverse ways to the well-being of others." More broadly, the Law School seeks "a mix of students with varying backgrounds and experiences who will respect and learn from each other." . . . The hallmark of [its] policy is its focus on academic ability coupled with a flexible assessment of applicants' talents, experiences, and potential "to contribute to the learning of those around them." The policy requires admissions officials to evaluate each applicant based on all the information available in the file, including a personal statement, letters of recommendation, and an essay describing the ways in which the applicant will contribute to the life and diversity of the Law School. . . .

The policy aspires to "achieve that diversity which has the potential to enrich everyone's education and thus make a law school class stronger than the sum of its parts." The policy does not restrict the types of diversity contributions eligible for "substantial weight" in the admissions process, but instead recognizes "many possible bases for diversity admissions." The policy does, however, reaffirm the Law School's long-standing commitment to "one particular type of diversity," that is, "racial and ethnic diversity with special reference to the inclusion of students from groups which have been historically discriminated against, like African-Americans, Hispanics, and Native Americans, who without this commitment might not be represented in our student body in meaningful numbers." . . .

---

539 U.S. 244 and 306 (2003).

II

. . . Today, we hold that the Law School has a compelling interest in attaining a diverse student body.

The Law School's educational judgment that such diversity is essential to its educational mission is one to which we defer. . . .

As part of its goal of "assembling a class that is both exceptionally academically qualified and broadly diverse," the Law School seeks to "enroll a 'critical mass' of minority students." The Law School's interest is not simply "to assure within its student body some specified percentage of a particular group merely because of its race or ethnic origin." That would amount to outright racial balancing, which is patently unconstitutional. Rather, the Law School's concept of critical mass is defined by reference to the educational benefits that diversity is designed to produce.

These benefits are substantial. As the District Court emphasized, the Law School's admissions policy promotes "cross-racial understanding," helps to break down racial stereotypes, and "enables [students] to better understand persons of different races." These benefits are "important and laudable," because "classroom discussion is livelier, more spirited, and simply more enlightening and interesting" when the students have "the greatest possible variety of backgrounds."

The Law School's claim of a compelling interest is further bolstered by its *amici*, who point to the educational benefits that flow from student body diversity. In addition to the expert studies and reports entered into evidence at trial, numerous studies show that student body diversity promotes learning outcomes, and "better prepares students for an increasingly diverse workforce and society, and better prepares them as professionals."

These benefits are not theoretical but real, as major American businesses have made clear that the skills needed in today's increasingly global marketplace can only be developed through exposure to widely diverse people, cultures, ideas, and viewpoints. What is more, high-ranking retired officers and civilian leaders of the United States military assert that, "[b]ased on [their] decades of experience," a "highly qualified, racially diverse officer corps . . . is essential to the military's ability to fulfill its principle mission to provide national security." The primary sources for the Nation's officer corps are the service academies and the Reverse Officers Training Corps (ROTC), the latter comprising students already admitted to participating colleges and universities. At present, "the military cannot achieve an officer corps that is *both* highly qualified *and* racially diverse unless the service academies and the ROTC used limited race-conscious recruiting and admissions policies." . . .

III

To be narrowly tailored, a race-conscious admissions program cannot use a quota system—it cannot "insulat[e] each category of applicants with certain desired qualifications from competition with all other applicants." Instead, a university may consider race or ethnicity only as a " 'plus' in a particular applicant's file," without "insulat[ing] the individual from comparison with all other candidates for the available seats." In other words, an admissions program must be "flexible enough to consider all pertinent elements of diversity in light of the particular qualifications of each applicant, and to place them on the same footing for consideration, although not necessarily according them the same weight." . . .

That a race-conscious admissions program does not operate as a quota does not, by itself, satisfy the requirement of individualized consideration. When using race as a "plus" factor in university admissions, a university's admissions program must remain flexible enough to ensure that each applicant is evaluated as an individual and not in a way that makes an applicant's race or ethnicity the defining feature of his or her application. The importance of this individualized consideration in the context of a race-conscious admissions program is paramount.

Here, the Law School engages in a highly individualized, holistic review of each applicant's file, giving serious consideration to all the ways an applicant might contribute to a diverse educational environment. The Law School affords this individualized consideration to applicants of all races. There is no policy, either *de jure* or *de facto*, of automatic acceptance or rejection based on any

single "soft" variable. Unlike the program at issue in *Gratz* v. *Bollinger*, the Law School awards no mechanical, predetermined diversity "bonuses" based on race or ethnicity. . . .

We also find that . . . the Law School's race conscious admissions program adequately ensures that all factors that may contribute to student body diversity are meaningfully considered alongside race in admissions decisions. With respect to the use of race itself, all underrepresented minority students admitted by the Law School have been deemed qualified. By virtue of our Nation's struggle with racial inequality, such students are both likely to have experiences of particular importance to the Law School's mission, and less likely to be admitted in meaningful numbers on criteria that ignore those experiences. . . .

What is more, the Law School actually gives substantial weight to diversity factors besides race. The Law School frequently accepts nonminority applicants with grades and test scores lower than underrepresented minority applicants (and other nonminority applicants) who are rejected. This shows that the Law School seriously weighs many other diversity factors besides race that can make a real and dispositive difference for nonminority applicants as well. . . .

We acknowledge that "there are serious problems of justice connected with the idea of preference itself." Narrow tailoring, therefore requires that a race-conscious admissions program not unduly harm members of any racial group. . . .

We are satisfied that the Law School's admissions program does not. Because the Law School considers "all pertinent elements of diversity," it can (and does) select nonminority applicants who have greater potential to enhance student body diversity over underrepresented minority applicants. . . .

We are mindful, however, that "[a] core purpose of the Fourteenth Amendment was to do away with all governmentally imposed discrimination based on race." Accordingly, race-conscious admissions policies must be limited in time. This requirement reflects that racial classifications, however compelling their goals, are potentially so dangerous that they may be employed no more broadly than the interest demands. Enshrining a permanent justification for racial preferences would offend this fundamental equal protection principle. . . . It has been 25 years since Justice Powell first approved the use of race to further an interest in student body diversity in the context of public higher education. Since that time, the number of minority applicants with high grades and test scores has indeed increased. We expect that 25 years from now, the use of racial preferences will no longer be necessary to further the interest approved today.

IV

In summary, the Equal Protection Clause does not prohibit the Law School's narrowly tailored use of race in admissions decisions to further a compelling interest in obtaining the educational benefits that flow from a diverse student body. . . . The judgment of the Court of Appeals for the Sixth Circuit, accordingly, is affirmed.

It is so ordered.

Justice Thomas, dissenting: Frederick Douglass, speaking to a group of abolitionists almost 140 years ago, delivered a message lost on today's majority:

> In regard to the colored people, there is always more that is benevolent, I perceive, than just, manifested towards us. What I ask for the negro is not benevolence, not pity, not sympathy, but simply *justice.* The American people have always been anxious to know what they shall do with us. . . . I have had but one answer from the beginning. Do nothing with us! Your doing with us has already played the mischief with us. Do nothing with us! . . . If the negro cannot stand on his own legs, let him fall also. All I ask is, give him a chance to stand on his own legs! Let him alone! . . . Your interference is doing him positive injury.

Like Douglass, I believe blacks can achieve in every avenue of American life without the meddling of university administrators. Because I wish to see all students succeed whatever their color, I share, in some respect, the sympathies of those who sponsor the type of discrimination advanced by the University of Michigan Law School (Law School). The Constitution does not, however, tolerate institutional devotion to the status quo in

admissions policies when such devotion ripens into racial discrimination. . . .

No one would argue that a university could set up a lower general admission standard and then impose heightened requirements only on black applicants. Similarly, a university may not maintain a high admission standard and grant exemptions to favored races. The Law School, of its own choosing, and for its own purposes, maintains an exclusionary admissions system that it knows produces racially disproportionate results. Racial discrimination is not a permissible solution to the self-inflicted wounds of this elitist admissions policy. . . .

I

One must also consider the Law School's refusal to entertain changes to its current admissions system that might produce the same educational benefits. The Law School adamantly disclaims any race-neutral alternative that would reduce "academic selectivity," which would in turn "require the Law School to become a very different institution, and to sacrifice a core part of its educational mission." In other words, the Law School seeks to improve marginally the education it offers without sacrificing too much of its exclusivity and elite status.

. . . The Court upholds the use of racial discrimination as a tool to advance the Law School's interest in offering a marginally superior education while maintaining an elite institution. . . . There is no pressing public necessity in maintaining a public law school at all and, it follows, certainly not an elite law school. Likewise, marginal improvements in legal education do not qualify as a compelling state interest. . . . Even if the Law School's racial tinkering produces tangible educational benefits, a marginal improvement in legal education cannot justify racial discrimination where the Law School has no compelling interest in either its existence or in its current educational and admissions policies. . . .

II

Putting aside the absence of any legal support for the majority's reflexive deference, there is much to be said for the view that the use of tests and other measures to "predict" academic performance is a poor substitute for a system that gives every applicant a chance to prove he can succeed in the study of law. The rallying cry that in the absence of racial discrimination in admissions there would be a true meritocracy ignores the fact that the entire process is poisoned by numerous exceptions to "merit." For example, in the national debate on racial discrimination in higher education admissions, much has been made of the fact that elite institutions utilize a so-called "legacy" preference to give the children of alumni an advantage in admissions. This, and other, exceptions to a "true" meritocracy give the lie to protestations that merit admissions are in fact the order of the day at the Nation's universities. The Equal Protection Clause does not, however, prohibit the use of unseemly legacy preferences or many other kinds of arbitrary admissions procedures. What the Equal Protection Clause does prohibit are classifications made on the basis of race.

. . . Since its inception, selective admissions has been the vehicle for racial, ethnic, and religious tinkering and experimentation by university administrators. The initial driving force for the relocation of the selective function from the high school to the universities was the same desire to select racial winners and losers that the Law School exhibits today. Columbia, Harvard, and others infamously determined that they had "too many" Jews, just as today the Law School argues it would have "too many" whites if it could not discriminate in its admissions process.

Columbia employed intelligence tests precisely because Jewish applicants, who were predominantly immigrants, scored worse on such tests. Thus, Columbia could claim (falsely) that "[w]e have not eliminated boys because they were Jews and do not propose to do so. We have honestly attempted to eliminate the lowest grade of applicant [through the use of intelligence testing] and it turns out that a good many of the low grade men are New York City Jews." In other words, the tests were adopted with full knowledge of their disparate impact.

Similarly no modern law school can claim ignorance of the poor performance of blacks, relatively speaking, on the Law School Admissions Test (LSAT). Nevertheless, law schools continue to use the test and then attempt to "correct" for black underperformance by using racial discrimination in

admissions so as to obtain their aesthetic student body. The Law School's continued adherence to measures it knows produce racially skewed results is not entitled to deference by this Court. . . .

Having decided to use the LSAT, the Law School must accept the constitutional burdens that come with this decision. The Law School may freely continue to employ the LSAT and other allegedly merit-based standards in whatever fashion it likes. What the Equal Protection Clause forbids, but the Court today allows, is the use of these standards hand-in-hand with racial discrimination. An infinite variety of admissions methods are available to the Law School. . . .

The Court will not even deign to make the Law School try other methods, however, preferring instead to grant a 25-year license to violate the Constitution. And the same Court that had the courage to order the desegregation of all public schools in the South now fears, on the basis of platitudes rather than principle, to force the Law School to abandon a decidedly imperfect admissions regime that provides the basis for racial discrimination.

III

. . . I believe what lies beneath the Court's decision today are the benighted notions that one can tell when racial discrimination benefits (rather than hurts) minority groups, and that racial discrimination is necessary to remedy general societal ills.

. . . I must contest the notion that the Law School's discrimination benefits those admitted as a result of it. . . . Nowhere in any of the filings in this Court is any evidence that the purported "beneficiaries" of this racial discrimination prove themselves by performing at (or even near) the same level as those students who receive no preferences. . . .

The Law School tantalizes unprepared students with the promise of a University of Michigan degree and all of the opportunities that it offers. These overmatched students take the bait, only to find that they cannot succeed in the cauldron of competition. And this mismatch crisis is not restricted to elite institutions. . . .

It is uncontested that each year, the Law School admits a handful of blacks who would be admitted in the absence of racial discrimination. Who can differentiate between those who belong and those who do not? The majority of blacks are admitted to the Law School because of discrimination, and because of this policy all are tarred as undeserving. This problem of stigma does not depend on determinacy as to whether those stigmatized are actually the "beneficiaries" of racial discrimination. When blacks take positions in the highest places of government, industry, or academia, it is an open question today whether their skin color played a part in their advancement. The question itself is the stigma—because either racial discrimination did play a role, in which case the person may be deemed "otherwise unqualified," or it did not, in which case asking the question itself unfairly marks those blacks who would succeed without discrimination. . . .

Gratz v. Bollinger

Justice Souter, dissenting: . . . . *Grutter* reaffirms the permissibility of individualized consideration of race to achieve a diversity of students, at least where race is not assigned a preordained value in all cases. On the other hand, Justice Powell's opinion in *Regents of Univ. of Cal. v. Bakke* rules out a racial quota or set-aside, in which race is the sole fact of eligibility for certain places in a class. Although the freshman admissions system here is subject to argument on the merits, I think it is closer to what *Grutter* approves than to what *Bakke* condemns, and should not be held unconstitutional on the current record.

. . . Membership in an underrepresented minority is given a weight of 20 points on the 150-point scale. On the face of things, however, this assignment of specific points does not set race apart from all other weighted considerations. Nonminority students may receive 20 points for athletic ability, socioeconomic disadvantage, attendance at a socioeconomically

disadvantaged or predominantly minority high school, or at the Provost's discretion; they may also receive 10 points for being residents of Michigan, 6 for residence in an underrepresented Michigan county, 5 for leadership and service, and so on.

The Court nonetheless finds fault with a scheme that "automatically" distributes 20 points to minority applicants because "[t]he only consideration that accompanies this distribution of points is a factual review of an application to determine whether an individual is a member of one of these minority groups." The objection goes to the use of points to quantify and compare characteristics, or to the number of points awarded due to race, but on either reading the objection is mistaken.

The very nature of a college's permissible practice of awarding value to racial diversity means that race must be considered in a way that increases some applicants' chances for admission. Since college admission is not left entirely to inarticulate intuition, it is hard to see what is inappropriate in assigning some stated value to a relevant characteristic, whether it be reasoning ability, writing style, running speed, or minority race. The college simply does by a numbered scale what the law school accomplishes in its "holistic review"; the distinction does not imply that applicants to the undergraduate college are denied individualized consideration or a fair chance to compete on the basis of all the various merits their applications may disclose.

Nor is it possible to say that the 20 points convert race into a decisive factor comparable to reserving minority places as in *Bakke.* Of course we can conceive of a point system in which the "plus" factor given to minority applicants would be so extreme as to guarantee every minority applicant a higher rank than every nonminority applicant in the university's admissions system. But . . . the present record obviously shows that nonminority applicants may achieve higher selection point totals than minority applicants owing to characteristics other than race. . . . It suffices for me, as it did for the District Court, that there are no *Bakke*-like set-asides and that consideration of an applicant's whole spectrum of ability is no more ruled out by giving 20 points for race than by giving the same points for athletic ability or socioeconomic disadvantage.

JUSTICE GINSBURG, dissenting:

I

Educational institutions, the Court acknowledges, are not barred from any and all consideration of race when making admissions decisions. But the Court once again maintains that the same standard of review controls judicial inspection of all official race classifications. This insistence on "consistency" would be fitting were our Nation free of the vestiges of rank discrimination long reinforced by law. But we are not far distant from an overtly discriminatory past, and the effects of centuries of law-sanctioned inequality remain painfully evident in our communities and schools.

In the wake "of a system of racial caste only recently ended," large disparities endure. Unemployment, poverty, and access to health care vary disproportionately by race. Neighborhoods and schools remain racially divided. African-American and Hispanic children are all too often educated in poverty-stricken and underperforming institutions. Adult African-Americans and Hispanics generally earn less than whites with equivalent levels of education. Equally credentialed job applicants receive different receptions depending on their race. Irrational prejudice is still encountered in real estate markets and consumer transactions. "Bias both conscious and unconscious, reflecting traditional and unexamined habits of thought, keeps up barriers that must come down if equal opportunity and nondiscrimination are ever genuinely to become this country's law and practice."

The Constitution instructs all who act for the government that they may not "deny to any person . . . the equal protection of the laws." In implementing this equality instruction, as I see it, government decision-makers may properly distinguish between policies of exclusion and inclusion. Actions designed to burden groups long denied full citizenship stature are not sensibly ranked with measures taken to hasten the day when entrenched discrimination and its after effects have been extirpated. . . . "To say that two centuries of struggle for the most basic of civil rights have been mostly about freedom from racial categorization rather than freedom from racial oppressio[n] is to trivialize the lives and deaths of those who have suffered under racism." . . .

Our jurisprudence ranks race a "suspect" category, "not because [race] is inevitably an impermissible classification, but because it is one which usually, to our national shame, has been drawn for the purpose of maintaining racial inequality." But where race is considered "for the purpose of achieving equality," no automatic proscription is in order. For, as insightfully explained, "[t]he Constitution is both color blind and color conscious. To avoid conflict with the equal protection clause, a classification that denies a benefit, causes harm, or imposes a burden must not be based on race. In that sense, the Constitution is color blind. But the Constitution is color conscious to prevent discrimination being perpetuated and to undo the effects of past discrimination." . . .

The mere assertion of a laudable governmental purpose, of course, should not immunize a race-conscious measure from careful judicial inspection. Close review is needed "to ferret out classifications in reality malign, but masquerading as benign," and to "ensure that preferences are not so large as to trammel unduly upon the opportunities of others or interfere too harshly with legitimate expectations of persons in once-preferred groups."

II

Examining in this light the admissions policy employed by the University of Michigan's College of Literature, Science, and the Arts (College), and for the reasons well stated by Justice Souter, I see no constitutional infirmity. . . . Like other top-ranking institutions, the College has many more applicants for admission than it can accommodate in an entering class. Every applicant admitted under the current plan . . . is qualified to attend the College. The racial and ethnic groups to which the College accords special consideration (African-Americans, Hispanics, and Native-Americans) historically have been relegated to inferior status by law and social practice; their members continue to experience class-based discrimination to this day. There is no suggestion that the College adopted its current policy in order to limit or decrease enrollment by any particular racial or ethnic group, and no seats are reserved on the basis of race. Nor has there been any demonstration that the College's program unduly constricts admissions opportunities for students who do not receive special consideration based on race.

The stain of generations of racial oppression is still visible in our society, and the determination to hasten its removal remains vital. One can reasonably anticipate, therefore, that colleges and universities will seek to maintain their minority enrollment—and the networks and opportunities thereby opened to minority graduates—whether or not they can do so in full candor through adoption of affirmative action plans of the kind here at issue. Without recourse to such plans, institutions of higher education may resort to camouflage. For example, schools may encourage applicants to write of their cultural traditions in the essays they submit, or to indicate whether English is their second language. Seeking to improve their chances for admission, applicants may highlight the minority group associations to which they belong, or the Hispanic surnames of their mothers or grandparents. In turn, teachers' recommendations may emphasize who a student is as much as what he or she has accomplished. If honesty is the best policy, surely Michigan's accurately described, fully disclosed College affirmative action program is preferable to achieving similar numbers through winks, nods, and disguises.

REVIEW AND DISCUSSION QUESTIONS

1. Review and assess the benefits of educational diversity as explained by Justice O'Connor. In your view, what constitutes "diversity"? How important is it in higher education, and why?
2. O'Connor writes that the law school's policy is not unfair to non-minorities. Do you agree? What about the University of Michigan's undergraduate admissions process? Explain your answers.
3. Why shouldn't affirmative action be permanent, according to O'Connor? Will affirmative action no longer be necessary 25 years from now, as the Court suggests? Assess the following argument: "If affirmative action will be wrong 25 years from now, then it's wrong now. We shouldn't suspend the Constitution for 25 years."

4. Justice Thomas suggests that at the root of the problem is the law school's effort to both be diverse and have a highly selective admissions process. Do you agree? What alternatives to affirmative action are open to the law school for increasing diversity?
5. Assess Justice Thomas's argument that there is no pressing public necessity for Michigan to have a public law school, especially an elite law school, and that therefore it can have no compelling interest in improving diversity in legal education.
6. Do you agree with Justice Thomas that affirmative action stigmatizes all blacks? If so, is this a sufficient reason for rejecting affirmative action? Are blacks better off being left alone, as he implies when he quotes Frederick Douglass ("Do nothing with us!")?
7. Why does Justice Souter believe there is nothing wrong with granting minority undergraduate applicants 20 points? Is there some significant difference between the law school's admissions process and the undergraduate admissions process?
8. Assess Justice Ginsburg's contention that there's a difference between using race to exclude and taking race into account in order to promote equality.
9. Justice Ginsburg maintains that Michigan's fully disclosed affirmative action program is better than what is likely to happen if such programs are deemed unconstitutional. Explain.

Gender Discrimination

Michael M. v. Sonoma County Superior Court

This case concerns equality between the sexes. A seventeen-year-old male was convicted under California law for the statutory rape of a sixteen-year-old female. Men, but not women, are liable under the statute, which makes it illegal to have sex with a female under eighteen, whether or not she consents. Justice Rehnquist, writing for a plurality of the Supreme Court, upheld the California law, maintaining that it does not discriminate against men or women. In a dissenting opinion, Justice Brennan argues that California has not shown that this law will reduce pregnancies more effectively than one that does not make any distinction based on sex.

Mr. Justice Rehnquist: [We] have not held that gender-based classifications are "inherently suspect" and thus we do not apply so-called "strict scrutiny" to those classifications. Our cases have held, however, that the traditional minimum rationality test takes on a somewhat "sharper focus" when gender-based classifications are challenged. In *Reed v. Reed,* for example, the Court stated that a gender-based classification will be upheld if it bears a "fair and substantial relationship" to legitimate state ends, while in *Craig v. Boren,* the Court restated the test to require the classification to bear a "substantial relationship" to "important governmental objectives."

Underlying these decisions is the principle that a legislature may not "make overbroad generalizations based on sex which are entirely unrelated to any differences between men and women or which demean the ability or social status of the affected class." But because the Equal Protection Clause does not "demand that a statute necessarily apply equally to all persons" or require "things which are different in fact to be treated in law as though they were the same," this Court has consistently upheld statutes where the gender classification is not invidious, but rather realistically reflects the fact that the sexes are not similarly situated in certain circumstances. . . .

---

450 U.S. 464 (1981). Some citations omitted.

Applying those principles to this case, the fact that the California Legislature criminalized the act of illicit sexual intercourse with a minor female is a sure indication of its intent or purpose to discourage that conduct. Precisely why the legislature desired that result is of course somewhat less clear. . . .

The justification for the statute offered by the State, and accepted by the Supreme Court of California, is that the legislature sought to prevent illegitimate teenage pregnancies. . . .

We are satisfied not only that the prevention of illegitimate pregnancy is at least one of the "purposes" of the statute, but also that the State has a strong interest in preventing such pregnancy. . . .

We need not be medical doctors to discern that young men and young women are not similarly situated with respect to the problems and the risks of sexual intercourse. Only women may become pregnant, and they suffer disproportionately the profound physical, emotional, and psychological consequences of sexual activity. The statute at issue here protects women from sexual intercourse at an age when those consequences are particularly severe.

[Although petitioner concedes that the State has a "compelling" interest in preventing teenage pregnancy, he contends that the "true" purpose of {the statute} is to protect the virtue and chastity of young women. As such, the statute is unjustifiable because it rests on archaic stereotypes. What we have said above is enough to dispose of that contention. The question for us—and the only question under the Federal Constitution—is whether the legislation violates the Equal Protection Clause of the Fourteenth Amendment, not whether its supporters may have endorsed it for reasons no longer generally accepted.]

The question thus boils down to whether a State may attack the problem of sexual intercourse and teenage pregnancy directly by prohibiting a male from having sexual intercourse with a minor female. We hold that such a statute is sufficiently related to the State's objectives to pass constitutional muster.

Because virtually all of the significant harmful and inescapably identifiable consequences of teenage pregnancy fall on the young female, a legislature acts well within its authority when it elects to punish only the participant who, by nature, suffers few of the consequences of his conduct. It is hardly unreasonable for a legislature acting to protect minor females to exclude them from punishment. Moreover, the risk of pregnancy itself constitutes a substantial deterrence to young females. No similar natural sanctions deter males. A criminal sanction imposed solely on males thus serves to roughly "equalize" the deterrents on the sexes. . . .

[We] cannot say that a gender-neutral statute would be as effective as the statute California has chosen to enact. The State persuasively contends that a gender-neutral statute would frustrate its interest in effective enforcement. Its view is that a female is surely less likely to report violations of the statute if she herself would be subject to criminal prosecution. . . .

There remains only petitioner's contention that the statute is unconstitutional as it is applied to him because he, like [his partner], was under 18 at the time of sexual intercourse. Petitioner argues that the statute is flawed because it presumes that as between two persons under 18, the male is the culpable aggressor. We find petitioner's contentions unpersuasive. Contrary to his assertions, the statute does not rest on the assumption that males are generally the aggressors. It is instead an attempt by a legislature to prevent illegitimate teenage pregnancy by providing an additional deterrent for men. The age of the man is irrelevant since young men are as capable as older men of inflicting the harm sought to be prevented.

Justice Brennan, dissenting: The State of California vigorously asserts that the "important governmental objective" to be served by [the statute] is the prevention of teenage pregnancy. It claims that its statute furthers this goal by deterring sexual activity by males—the class of persons it considers more responsible for causing those pregnancies. [In a remarkable display of sexual stereotyping, the California Supreme Court stated: "The Legislature is well within its power in imposing criminal sanctions against males, alone, because they are the *only* persons who may physiologically cause the result which the law properly seeks to avoid."] But even assuming that prevention of teenage pregnancy is an important governmental objective and that it is in fact an objective of [the statute], California still has the burden of proving

that there are fewer teenage pregnancies under its gender-based statutory rape law than there would be if the law were gender neutral. To meet this burden, the State must show that because its statutory rape law punishes only males, and not females, it more effectively deters minor females from having sexual intercourse. . . .

A State's bare assertion that its gender-based statutory classification substantially furthers an important governmental interest is not enough to meet its burden of proof. . . . Rather, the State must produce evidence that will persuade the Court that its assertion is true. [Even] assuming that a gender-neutral statute would be more difficult to enforce, the State has still not shown that those enforcement problems would make such a statute less effective than a gender-based statute in deterring minor females from engaging in sexual intercourse. Common sense, however, suggests that a gender-neutral statutory rape law is potentially a *greater* deterrent of sexual activity than a gender-based law, for the simple reason that a gender-neutral law subjects both men and women to criminal sanctions and thus arguably has a deterrent effect on twice as many potential violators. Even if fewer persons were prosecuted under the gender-neutral law, as the State suggests, it would still be true that twice as many persons would be *subject* to arrest.

Until very recently, no California court or commentator had suggested that the purpose of California's statutory rape law was to protect young women from the risk of pregnancy. Indeed, the historical development of [the statute] demonstrates that the law was initially enacted on the premise that young women, in contrast to young men, were to be deemed legally incapable of consenting to an act of sexual intercourse. Because their chastity was considered particularly precious, those young women were felt to be uniquely in need of the State's protection. In contrast, young men were assumed to be capable of making such decisions for themselves; the law therefore did not offer them any special protection. . . .

I would hold that [the statute] violates the Equal Protection Clause of the Fourteenth Amendment. . . .

REVIEW AND DISCUSSION QUESTIONS

1. Explain the basic legal question before the Court in this case.
2. What are Justice Rehnquist's reasons for upholding the law?
3. Explain Justice Brennaa's reasons for dissenting.
4. In your view, does the California law discriminate against men, against women, against neither, or against both?
5. Why do you think California passed this law? Could a gender-neutral law have served its purposes equally well?
6. Should the state have any laws regulating or prohibiting sex between minors?

The Bill of Rights and the Fourteenth Amendment

AMENDMENT I

Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the Government for a redress of grievances.

AMENDMENT II

A well regulated Militia, being necessary to the security of a free State, the right of the people to keep and bear Arms, shall not be infringed.

AMENDMENT III

No Soldier shall, in time of peace be quartered in any house, without the consent of the Owner, nor in time of war, but in a manner to be prescribed by law.

AMENDMENT IV

The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.

AMENDMENT V

No person shall be held to answer for a capital, or otherwise infamous crime, unless on a presentment of indictment of a Grand Jury, except in cases arising in the land or naval forces, or in the Militia, when in actual service in time of War or public danger; nor shall any person be subject for the same offence to be twice put in jeopardy of life or limb; nor shall be compelled in any criminal case to be a witness against himself; nor be deprived of life, liberty, or property, without due process of law, nor shall private property be taken for public use, without just compensation.

AMENDMENT VI

In all criminal prosecutions, the accused shall enjoy the right to a speedy and public trial, by an impartial jury of the State and district wherein the crime shall have been committed, which district shall have been previously ascertained by law, and to be informed of the nature and cause of the accusation; to be confronted with the witnesses against him; to

have compulsory process for obtaining witnesses in his favor, and to have the Assistance of Counsel for his defence.

AMENDMENT VII

In suits at common law, where the value in controversy shall exceed twenty dollars, the right of trial by jury shall be preserved, and no fact tried by a jury, shall be otherwise re-examined in any Court of the United States, than according to the rules of the common law.

AMENDMENT VIII

Excessive bail shall not be required, nor excessive fines imposed, nor cruel and unusual punishment inflicted.

AMENDMENT IX

The enumeration in the Constitution, of certain rights, shall not be construed to deny or disparage others retained by the people.

AMENDMENT X

The powers not delegated to the United States by the Constitution, nor prohibited by it to the States, are reserved to the States respectively, or to the people.

AMENDMENT XIV

Section 1. All persons born or naturalized in the United States, and subject to the jurisdiction thereof, are citizens of the United States and of the State wherein they reside. No State shall make or enforce any law which shall abridge the privileges or immunities of citizens of the United States; nor shall any State deprive any person of life, liberty, or property, without due process of law; nor deny to any person within its jurisdiction the equal protection of the laws. . . .

ACKNOWLEDGMENTS

Lawyers' Ethics in an Adversary System: From Monroe H. Freedman, *Lawyers' Ethics in an Adversary System* (Indianapolis: Bobbs-Merrill Company, 1975). Reprinted by permission of the author. Some footnotes omitted.

An Alternative to the Adversary System: From John H. Langbein, "The German Advantage in Civil Procedure," *University of Chicago Law Review,* Vol. 52 (1985), p. 823.

Pure Legal Advocates and Moral Agents: Two Concepts of a Lawyer in an Adversary System: From Elliot D. Cohen, "Pure Legal Advocates and Moral Agents: Two Concepts of a Lawyer in an Adversary System," *Criminal Justice Ethics* 4 (Winter–Spring 1985). Reprinted by permission.

Rule of Law and the English Constitution: From A. V. Dicey, *Introduction to the Study of the Law of the Constitution*, 8th ed. (1914).

Eight Ways to Fail to Make Law: From Lon L. Fuller, *The Morality of Law* (New Haven: Yale University Press, 1964). Copyright © 1964 by Yale University Press. Reprinted by permission.

The Rule of Law and Its Virtues: From Joseph Raz © 1979, "The Rule of Law and Its Virtue." Reprinted with permission of the author from *The Authority of Law: Essays on Law and Morality* by Joseph Raz, Oxford University Press, 1979.

Grudge Informers and the Rule of Law: From H. L. A. Hart, "Positivism and the Separation of Law and Morals," 71 *Harvard Law Review* 90, 616–20 (1958). Copyright 1958 by the Harvard Law Review. Excerpts reprinted by permission of the Harvard Law Review Association.

The Problem of the Grudge Informer: From Lon L. Fuller, *The Morality of Law* (New Haven: Yale University Press, 1964). Copyright © 1964 by Yale University Press. Reprinted by permission.

Crito: From Plato, *The Dialogues of Plato,* 3d ed., trans. Benjamin Jowett (London: Oxford University Press, 1892).

The Justification of Civil Disobedience: Originally presented at the meetings of the American Political Science Association, September 1966. Some revisions have been made and two paragraphs have been added to the last section. Copyright © 1968 by John Rawls. Reprinted by permission.

On Not Prosecuting Civil Disobedience: From Ronald Dworkin, *Taking Rights Seriously* (Cambridge, MA: Harvard University Press, 1977), pp. 206–22. Reprinted by permission.

On Interpretation: The Adultery Clause of the Ten Commandments: From Sanford Levinson, "On Interpretation: The Adultery Clause of the Ten Commandments," *Southern California Law Review* 58, no. 2 (1985). Reprinted with the permission of the *Southern California Law Review.* The fictitious dates in the text have been altered.

Interpreting the Small Bird Act: From H. Pomerantz and S. Breslin, *Criminal Law Quarterly,* 137 (1965). Reprinted by permission of Canada Law Book, Inc.

A Case Study in Interpretation: The Mann Act: From Edward H. Levi, *An Introduction to Legal Reasoning* (Chicago: University of Chicago Press, 1948). Copyright © 1948 University of Chicago Press. Reprinted by permission.

Interpretation: From Tony Honoré, *About Law* (Oxford: Clarendon Press, 1995). Reprinted by permission of Oxford University Press.

Stare Decisis: The Use of Precedent: From C. Gordon Post, "Stare Decisis: The Use of Precedent," in *An Introduction to the Law* (Englewood Cliffs, NJ: Prentice Hall, 1963). Copyright © 1963 by Prentice Hall. Some footnotes omitted.

Constrained by Precedent: From Larry Alexander, "Constrained by Precedent," *Southern California Law Review* 63 (1989). Reprinted by permission.

"Vague" Constitutional Clauses: From Ronald Dworkin, *Taking Rights Seriously* (Cambridge, MA: Harvard University Press). Copyright © 1978 by Ronald Dworkin. Reprinted by permission of the author.

Interpreting the Constitution: From Antonin Scalia, *A Matter of Interpretation: Federal Courts and the Law.* Copyright © 1997 Princeton University Press. Reprinted by permission.

A Response to Scalia: From Ronald Dworkin, "Comment," in Antonin Scalia, *A Matter of Interpretation: Federal Courts and the Law.* Copyright © 1997 Princeton University Press. Reprinted by permission.

In Defense of Looseness: The Supreme Court and Gun Control: From Richard A. Posner, "In Defense of Looseness," *The New Republic*, August 27, 2008. Reprinted by permission.

Traditional Natural Law Theory: From Dennis Patterson, ed., *A Companion to Philosophy of Law and Legal Theory* (Oxford: Blackwell Publishers, Ltd., 1996). Reprinted by permission.

Summa Theologica: Selections from *Summa Theologica,* reprinted from Anton C. Pegis, ed., *The Basic Writings of Saint Thomas Aquinas* (New York: Random House, 1945); used by permission of the A. C. Pegis Estate. Footnotes omitted.

Blackstone's Commentaries: From William Blackstone, *Commentaries on the Laws of England* (1765–69), Vol. 1.

The Province of Jurisprudence Determined: From John Austin, "The Province of Jurisprudence Determined," *Lectures on Jurisprudence,* 5th ed.

The System of Law: From Joseph H. Beale, *A Treatise on the Conflict of Laws* (1935).

The Path of the Law: From Oliver Wendell Holmes, "The Path of the Law," *Harvard Law Review,* Vol. 10 (1897), pp. 457–68.

Realism and the Law: From Jerome Frank, *Law and the Modern Mind* (New York: Doubleday and Co., 1963). Copyright 1930, 1933, 1949 by Coward McCann, Inc., and copyright 1930 by Brentano's Inc., from Anchor Books Edition, 1963. Copyright renewed in 1958 by Florence K. Frank. Reprinted by arrangement with the estate of Barbara Frank Kristein.

Positivism and the Separation of Law and Morals: From H. L. A. Hart, "Positivism and the Separation of Law and Morals," 71 *Harvard Law Review* 593 (1958). Copyright © 1958 by the Harvard Law Review Association. Reprinted by permission.

Law as the Union of Primary and Secondary Rules: © Oxford University Press 1961, 1994. Reprinted from *The Concept of Law* by H. L. A. Hart (2nd edition, 1994) by permission of Oxford University Press.

The Model of Rules: From Ronald Dworkin, *Taking Rights Seriously* (Cambridge, MA: Harvard University Press, 1977). Reprinted by permission. Some notes omitted.

Natural Law Revisited: From Ronald A. Dworkin, "'Natural' Law Revisited," *University of Florida Law Review* 34 (1982). Reprinted by permission of the *University of Florida Law Review.*

The Dependence of Morality on Law: © Oxford University Press 1993. Reprinted from *The Oxford Journal of Legal Studies,* Vol. 13, no. 1 (1993) by permission of Oxford University Press.

International Law: © Oxford University Press 1961, 1994. Reprinted from *The Concept of Law* by H.L.A. Hart (2nd edition, 1994) by permission of Oxford University Press.

The Economic Approach to Law: Reprinted by permission of the publisher form Richard A. Posner, *The Problems of Jurisprudence* (Cambridge, Mass: Harvard University Press. Copyright © 1990 by the President and Fellows of Harvard College.)

Law and Economics: An Analysis and Critique: From *Arguing About Law,* by A. Altman. © 1996. Reprinted with permission of Wadsworth-Cengage Learning.

Who Should Be Punished? From *The Case of the Dog Provetie.* Reported in *The South African Law Journal,* Vol. 24 (1907), pp. 232–4.

The Utilitarian Theory of Criminal Punishment: From Richard B. Brandt, *Ethical Theory* (Englewood Cliffs, NJ: Prentice-Hall, 1959). Reprinted by permission of the author.

Persons and Punishment: From Herbert Morris, "Persons and Punishment," *The Monist* 52, no. 4 (October 1968). Copyright © 1968 by *The Monist,* La Salle, Illinois. Reprinted by permission of *The Monist.*

The Moral Education Theory of Punishment: From Jean Hampton, "The Moral Education Theory of Punishment," *Philosophy & Public Affairs* 13, no. 3 (1984). Copyright © 1986 by Princeton University Press. Reprinted by permission of Blackwell Publishing.

Restitution: A New Paradigm of Criminal Justice: From Randy E. Barnett, "Restitution: A New Paradigm of Criminal Justice," *Ethics* 87, no. 4 (July 1977). Copyright © 1977 by the University of Chicago Press. Reprinted by permission of the University of Chicago Press. Some footnotes omitted.

Criminal Justice and the Negotiated Plea: From Kenneth Kipnis, "Criminal Justice and the Negotiated Plea," *Ethics* 36 (1976). Copyright © 1976 by the University of Chicago Press. Reprinted by permission of the author and the publisher. Some footnotes omitted.

Convicting the Innocent: From James McCloskey, "Convicting the Innocent," *Criminal Justice Ethics* 8, no. 1 (winter–spring 1989). Copyright © 1989 by *Criminal Justice Ethics,* 899 Tenth Avenue, New York, NY 10019. Reprinted by permission.

The Dirty Harry Problem: Reprinted from *The Annals,* vol. 452 (November 1980) by permission of Sage Publications. Copyright © 1980 by Sage Publications. Some section headings modified.

A Debate on the Exclusionary Rule: From Malcolm Richard Wilkey and Stephen H. Sachs, "The Exclusionary Rule: A Prosecutor's Defense," *Criminal Justice Ethics* 1, no. 2 (summer-fall 1982). Reprinted by permission of the author and *Criminal Justice Ethics.* Some footnotes omitted.

Electric Shock: The Fairest Punishment of All: From Graeme Newman, *Just and Painful: A Case for the Corporal Punishment of Criminals,* 2nd ed. (Albany: Harrow & Heston, 1983). Reprinted by permission of the Criminal Justice Press.

Survival on a Lifeboat: From *Law Reports, Queen's Bench Division,* Vol. 14 (1884–85) (London: William Clowes and Sons, 1885).

The Principles of Criminal Law: From Richard B. Brandt, *Ethical Theory* (Englewood Cliffs, NJ: Prentice-Hall, 1959). Reprinted by permission of the author.

Intention: From H. L. A. Hart, *Punishment and Responsibility* (New York: Oxford University Press, 1969). Copyright © 1969 by Oxford University Press. Reprinted by permission.

Attempting the Impossible: The Crime That Never Was: From Leo Katz, "The Crime that Never Was: A Fake Opinion in a Fake Case Involving Fakes," *Bad Acts and Guilty Minds* (Chicago: University of Chicago Press, 1987, pp. 276–93). Copyright © 1987 by the University of Chicago. All Rights Reserved. Reprinted with permission of the University of Chicago.

Rape, Force, and Consent: From Susan Estrich, "Rape," *Yale Law Journal* 95, no. 6 (May 1986). Reprinted by permission of the Yale Law Journal Company and Fred B. Rothman & Company.

Is the Insanity Test Insane?: From R. J. Gerber, "Is the Insanity Test Insane?" *The Journal of American Jurisprudence* 20 (1975). Reprinted by permission. Some footnotes omitted.

What Is So Special About Mental Illness?: From Joel Feinberg, *Doing and Deserving: Essays in the Theory of Responsibility* (Princeton: Princeton University Press, 1970). Copyright © 1970 by Princeton University Press. Reprinted by permission of Princeton University Press.

The Efficiency of the Common Law: From *Arguing About Law,* by A. Altman. © 1996. Reprinted with permission of Wadsworth-Cengage Learning.

Negligence: From William Prosser, *The Law of Torts,* 4th ed. (West Publishing Company, 1971). Reprinted by permission of the West Publishing Company.

Tort Liability and Corrective Justice: Reprinted by permission from Jeffrie G. Murphy and Jules L. Coleman, *The Philosophy of Law,* rev. ed. Copyright © 1990 by Westview Press. Reprinted by permission of Westview Press, a member of Perseus Books, L. L. C.

Loss, Agency, and Responsibility for Outcomes: From Steven Perry, "Loss, Agency, and Responsibility for Outcomes," in *Tort Theory,* ed. K. Cooper-Stephenson and E. Gibson (North York, Ontario: Captus Press, 1993). Reprinted by permission.

Property: From John Locke, *Second Treatise of Government: An Essay Concerning the Origin, Extent and End of Civil Government* (1690).

Property and Sovereignty: From Morris Raphael Cohen, "Property and Sovereignty," 13 *Cornell Law Quarterly* 8 (1927). Copyright © 1927 by Cornell University. All rights reserved. Reprinted by permission.

Intellectual Property: Reprinted by permission from *Daedalus,* Vol. 131, no. 2 (Spring 2002). Copyright © 2002 American Academy of Arts and Sciences.

The Basis of Contract: From Morris Raphael Cohen, "The Basis of Contract," 46 *Harvard Law Review* 553 (1933). Copyright © 1933 by Harvard Law Review Association. Reprinted by permission.

Contract as Promise: From Charles Fried, *Contract as Promise: A Theory of Contractual Obligation* (Cambridge, MA: Harvard University Press, 1981). Copyright © 1981 by the President and Fellows of Harvard College. Reprinted by permission. Notes omitted.

A Letter Concerning Toleration: From John Locke, "A Letter Concerning Toleration," in *John Locke on Politics and Education,* ed. Howard R. Penniman (New York: Van Norstrand Company, 1947).

On Liberty: From John Stuart Mill, *On Liberty* (1859).

Judicial Activism and Gay Marriage: A Debate: From *Harvard Law Bulletin* (Harvard Law School, Summer 2004), p. 8.

Of the Liberty of Thought and Discussion: From John Stuart Mill, *On Liberty* (1859).

Racial Equality and Affirmative Action: From Ronald Dworkin, "Why Bakke Has No Case," *New York Review of Books,* November, 10, 1977. Copyright © 1977, Nyrev, Inc. Reprinted by permission of *The New York Review of Books.*

ARTHUR 0205708099
READINGS IN PHILOSOPHY OF LAW
A

MySearchLab
to master a writi

- ✓ Round-the-clock access to reliable content for internet research projects includes thousands of full articles from the EBSCO ContentSelect database, census data from Social Explorer™, daily news feeds from The Associated Press, and primary and secondary source documents from the Pearson bookshelf.
- ✓ Step-by-step tutorials present complete overviews of the research and writing process.
- ✓ Pearson SourceCheck™ offers an easy way to detect accidental plagiarism issues, and our exclusive tutorials teach how to avoid them in the future.
- ✓ AutoCite helps to correctly cite sources in a variety of formats.
- ✓ Our exclusive online handbook provides grammar and usage support.

To order this book with MySearchLab access at no extra charge, use ISBN: 0-205-79351-7

PEARSON mysearchlab™

www.mysearchlab.com

WWW.SBSPSU.COM
ARTHUR
READINGS IN PHILOSOPHY OF LAW
5TH 10 PH
978-0-205-70809-3
2020-01-29
8491842-8
10 A
USED

Prentice Hall
is an imprint of
PEARSON
www.pearsonhig.